



SEQUENCE LISTING

<110> ZHOU, PENGBO
HOWLEY, PETER M.

<120> TARGETED PROTEOLYSIS BY RECRUITMENT TO UBIQUITIN
PROTEIN LIGASES

<130> HMV-043.01

<140> 09/415,795

<141> 1999-10-08

<150> 60/103,787

<151> 1998-10-09

<160> 49

<170> PatentIn Ver. 2.1

<210> 1

<211> 20383

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 1

```
ggatccgtgt cgtaactgcg ttccgtacac cttaaaatga actttccagc aggtgcacca 60
tcttccaatc cattgaaaaa tcttactgat tcttcgacct cctcctggct taagtttgct 120
tccaaatgaa tatatagcag ccaaataaat atgtttgatc tatcctttgg ctcgattttc 180
gaatagtcga gtatatcaaa taatgttgtt gcataagctg cttttgattc ttgtttttca 240
ttttccgctt cegtttcttt tacttcattc tgttcattca ccatactctg ctccctgagct 300
ttcgattcga cttccacata cttggaaaga aacctcttgt caattaggcc ggaattatca 360
catatgctaa acaacatatt aataatgtta aatttggtta cgggtgtccac ttcaccatca 420
ttttggctga cctttttgta agtctccatc aatgagagtt gctcattgcc ccacggcaga 480
tttttcagaa tagatttcag tctagggtga tcttgcaaaa gtttcggatc attggcatgt 540
aattgtaaaa ctggcacgga gtgaaaagtt ctcaattgcy atgtcatttc taaataaaaa 600
tttatagtgg tgttcaacct accaatgttg accaataacg ctaaaacaca agtagaaaaa 660
gctacttggt ggtctaataa tagtttatct ctgagaattt tggaacattt gatagaagt 720
gctattgtca acacatacaa ttgtgcgaac gtcaatttat cgttaaatac aaatgtcctt 780
gcacgtact cttgtcggg cactgtgatt gataatttat cgccatcgat attaatgggt 840
acaatagagt tctgtatag gggttttaga acattagtga agatttgcct cttgtccatc 900
aacagttcct gcataaaact gaactgaatg tcttttctat tgaatggctc gccgtcaggt 960
ttcttcaaat gacgattgta ttactggca ttgctttctt tgggtttcct tctaacagt 1020
ttagttgtca ctggcccggt tgcgtaact gctcctgttg cttcaggcac tgtagaaggt 1080
tcgaagttga gcaagatgtt ggtattttct tcttcacttt gcgtggagtc atcatcattt 1140
ttatcgctga tttcaatctc gtgtataact ggaatgcttt ttgtagtcac agtctcggtc 1200
tctgattgaa tggagcttgg ggatgcttct tcttgtgatc ccgattttac agattgcatt 1260
ggagaatctg cttttgtttc gtcagaattg tcttctcgta gttggaaagt atcatgaatt 1320
ggatcgata ctctcttccc catctttaat tagagtgtcg cctttattta gctttcgatt 1380
ttacttagtt aatcatgaac tgtttccaat accatatcag cattcattgg gcgttcttta 1440
cttacttgta cegttaccgc attcgagggt aacctgtttt tcgttgatca cgatatgcta 1500
cagaattgtc ttaatatggg ctaagaaaaa aaaaagctctg attatttctg atactgcaaa 1560
atatatactg gcttggttaa gaaaagtggt gctttagttc ctttacatca agagtcgtta 1620
aattgttttg tgcataaca gtatgtgtct tttgagacat caatcgactc tcaggttttc 1680
cttctgttca tcttctgtt ttcggtaaat gtcagccagc atatttcaat tctctctttt 1740
ttgaaataac attatgtaat ctttcaatcc tttctatcca acgtctgcta aataatttat 1800
tcaaatgaca tgtagaccag aaaatcgaa gttggataat ttttcatcct gaatcatgtt 1860
ggctctcaact agagcaattt tttcacgtcc tctgacacaa aactcgataa actgtcgcat 1920
cagtttaaaa aaatatgtac gtgtataata actgaattta aaaggatagt aaataatttt 1980
tgcattttat ggaatgtcaa aatactaata agtcgcaaat atagctatca aataccataa 2040
```

tttagctact	tatagaaaga	tgcccaaate	ccgacaaaa	aggaccattg	cgctcttcctc	2100
gtcagttttt	tatggaagtt	cacctttcca	aatgatggc	tacatcaaag	taatggaact	2160
cgtatcacac	attgtcattg	aaataaatca	ttcacctacc	gcaacaacgg	atgaaacgag	2220
aaagcagaat	aatccggagc	tgaaagtga	agaaccagtt	tgtaacctca	agaagtggga	2280
aaataacact	aacttttat	tggaagatca	tacgaagaat	aaaacaaaac	tttctagcac	2340
agataggata	cgtaagtgg	ttagaagaca	tatattaaag	gaagagatcg	aaatcctttc	2400
ccatggaaaa	caattgagta	gtattgatga	ggattattgc	ccttcaaagt	ttctttagg	2460
atgttcaaga	gatctaaata	aactcagatc	atttcaaaat	tttttaggaat	ataaaatttc	2520
tcagtaagaa	aaagaagatg	aatgcgttta	tagtatttta	agttgttcaa	tagaatatct	2580
tcaatatgtt	catatgtgga	ctaaaagaag	attttggtat	atgtctataa	atttattgtt	2640
ggtaaacgcc	attaagcagc	tttccaagaa	tatgttatca	ttgaggagaa	gtacgaaaac	2700
tgattatgca	cgattaggaa	aaattctgca	aggagtatta	aaaaatggcc	ttcgagaaga	2760
aatgaaatct	ccttcgatga	gcaatggctc	tatggataat	ttgatagaca	tatgaaaaag	2820
tcaaaagtgt	ttttggccat	atgcagtaaa	aacttaccct	tcaaattgca	aaaccattag	2880
ttttattata	tttttttttt	tttctttcat	tgaagtatac	acgaaatccc	atacgcaaat	2940
aaacagtcgt	tgtatcatca	atttcggctg	tctggatggc	gactacgaga	ggaagcttgc	3000
ttgccatgat	aggaccaatc	tcaattttta	atttttcatt	aactagctaa	tttctataaa	3060
tgctctaagt	ttcctttttc	gttttgtttc	tttatagtga	tgatgctgtt	attgtgacgc	3120
tatcaaatct	tatactgtca	tcaatcctca	tctcatcggc	tccggaacgc	gttgaatatg	3180
ttgttaaaca	agctctttat	agctagctat	agtagaaaag	gaaggaaaga	agtctaactcg	3240
ctgtattatc	gactgacaag	ggttgtcagt	tgccctcttt	tattttttgt	actcccctga	3300
atcgagcgga	tgtatttttt	aacaaagtat	tcagaattgt	ttttgtataa	tcaaaatttg	3360
tcgacgctaa	gaaaatgacg	gacgtcgtaa	ttccgctact	aggttttgtg	tgctcatcga	3420
tatttgacac	ctatttttca	cacttttaag	gttatatact	ttgtttcgga	tgtggggtac	3480
actcagttga	ctgattatga	aagatatggc	gtagccgttg	ggtaggggtt	attacacaga	3540
aatgaattat	accagtgaag	ttattctctt	tatgtttgtt	cctacatgtt	ttccaatcag	3600
aaggaattgt	tatataagga	cacaagaaat	gtgttggtac	gctctgcgcc	tcaatgacgg	3660
tttatctatc	taaaactatt	tctttctcgt	aataaaaaata	caataaataa	ttaataataa	3720
tatagttgat	cgatataaca	aaaaaataac	agcgaggatg	gttagttcaa	gtgtttccat	3780
tttggggact	agcgccaagg	catccacttc	tctaagtaga	aaggatgaaa	ttaaactaac	3840
ccctgaaaca	aggaagcta	gcttgacat	tccatacaaa	cccattattg	catactggac	3900
gggtgatggg	ctctgtctga	tgattgcctt	tggtggattc	atttttgggt	gggatacagg	3960
aaccatttca	gggtttatta	accaaacaga	tttcaagaga	aggtttgggt	agttacaaag	4020
ggacggcagt	tttcaactat	cagatgtcag	gacagggcta	attgtcggta	tcttcaacat	4080
aggttgtgct	ttaggtggcc	taacgctggg	acgcctgggc	gatatttatg	ggcgtaaaat	4140
cggcttaagt	tgtgttatac	tggtgtatgt	tggttggtatc	gtgatccaga	ttgcttcttc	4200
tgacaaatgg	tatcaatatt	ttattggtag	aattgtttct	ggaatgggtg	ttggagggtg	4260
tgctgtgctg	tcgccaactt	tgatctcaga	aatttcccca	aagcacctaa	gaggcacttg	4320
tgtctctttt	taccagctaa	tgattaccct	tggaattttc	ttgggctact	gtaccaatta	4380
tggtacaaag	aatatttcaa	attcaataca	gtggcgggtt	cccttgggtt	tgtgttttgc	4440
gtgggcaatc	tttatgggtga	ttggaatgg	tatggttccg	gaatcgccca	gatatttagt	4500
agaaaaaggt	aagtagaag	aagctagaag	gtccttggcc	aatcaaaaca	aggtcacagt	4560
tactgatcca	ggcgttgttt	ttgagtttga	tactatagtt	gcaaatatgg	aattagaaag	4620
ggctgttggg	aatgccagtt	ggcacgaact	cttctcaa	aaaggagcaa	ttctaccaag	4680
ggtaataatg	ggaatcggtt	tccagtcact	gcaacagctt	actggctgta	attatttttt	4740
ctactacggc	acgaccattt	tcaatgctgt	tggaatgcaa	gactctttcg	agacttccat	4800
tgtccttggg	gctgttaatt	ttgcttctac	atttgttgca	ctatacatgt	tggataaatt	4860
tgggcgctga	aaatgtttat	tgtgggggtc	tgccctcgatg	gcaatttgtt	togtcatatt	4920
cgccaccgtt	ggcgtcacta	gattatggcc	acaagggaag	gaccaacctt	cttcgcaaag	4980
tgctggtaat	gttatgatcg	tttttacttg	tttcttcatt	ttctcttttg	ccattacttg	5040
ggctcctatc	gcctatgtca	ttgtggcaga	aacttatcca	ttaagagtta	aaaatcgtgc	5100
catggccatt	gcggttgggt	cgaactggat	gtgggggttc	ttgattggat	ttttcacacc	5160
ctttatcact	agatccatag	gattttctta	tggtatgtt	ttcatgggtt	gcttaatctt	5220
ttcgtaactc	tacgttttct	tctttgtttg	cgaacaaag	ggattaactc	tggggaagt	5280
taatgaaatg	tacgaagaaa	gaataaagcc	atggaagtcc	ggaggttggg	ttccagttc	5340
tagaagaaca	ccacaaccaa	caagcagtac	accattagtt	attgttgata	gtaataaatt	5400
tctaaatatt	cttgactctt	ggtaaacaga	aataacaaca	gataatggat	tgatgcttta	5460
cttctatttc	atggagattg	gttttatata	atagccttta	ttaatggcgt	cacaaattga	5520
aaaaaaaaat	aaaaaaaaata	agacggacac	atgtgcgacg	ccggtgaata	aatgcatata	5580

agtagtttat	aagctagcta	ctactcagaa	ataatttcaa	aaaaaaaaag	aagtggggcac	5640
tttaaaatga	agattagcta	tagtaaatga	ctgtagtata	acaactacct	gctatctttc	5700
tgaaaaaatc	aggcaatgta	aggtaaaact	tgtaccatcg	acatatataa	tgttttgaga	5760
tataagtaac	tagagaccag	tttatacagg	atcttacctc	tttttaccgt	tatgaaagct	5820
ttattactgc	gttgtagta	aacacacata	ttattttcgg	gtagtctctg	cgatgttcca	5880
aggctttgca	tttatcattg	tttccgcttc	atcgggtgta	ttagttagt	tatctgtatg	5940
tgaagtatgt	ggataggtgc	ttctattatt	ggaacaaaac	accttaaacg	cgcactaggt	6000
tataggaag	ggtccattat	tcaaagaacg	gcttattgaa	aagtatgttg	taaagctcgg	6060
tcatatcgcg	catagccacg	ataaggcgtg	gtgctgcttc	atagattgag	aggcgagtc	6120
aatacacaaac	cataccggta	gtttgataaa	tgatactatc	attccggaag	ctctctagta	6180
agctgtaggt	ggcatatacg	gtatctatca	tctagtaata	gctattgatt	tttttctcga	6240
ctatgctata	tgtgttgta	tggttcactt	agaagtgaag	aagtagcaat	aaattagaca	6300
gacagaagta	attgaagcta	cattcaacaa	catggctaag	taaaaaagcg	tagaatatca	6360
tgatgggttt	ttttatcttt	taattgcttt	ctatgtagta	taatggaccg	atctttgcaa	6420
gtatatatct	gtatgtatcc	atatttagat	ggcagcaagc	aatatagatt	tgatgagctt	6480
atatcatctt	atcgctcttg	tccaaaaagt	cttgataaca	ttaaaagtca	ctaccgtcaa	6540
atccatcatc	aaatccgccc	tgaacccac	cagcatcatc	aaatccgccc	tccgacccac	6600
cagcatcatc	accgtagtaa	ttgttctcga	caacgacagt	gtctgggtcc	tcatagttgt	6660
ggtcgtaaaa	tgcgtgttct	agcatagctc	cacctaaaca	acccgcacca	acaccaagta	6720
aaccacccat	catggcaccc	cgtgtctcct	tacctttact	ggtgcttgcc	gcagttccat	6780
agtatgcttg	ctgtgggtct	gctgctggca	taggggcttg	ttgaggatag	taacgctgct	6840
gttgagggtg	ttgagggtac	tgaggctggt	gaggctgata	gtaacgaggt	tgctgtgctt	6900
gggccccctg	ttgcacttgc	ggtgttgatt	gagaagaata	aggtggagga	gcctgggtctg	6960
cctgttgacg	agaactcttc	tcatgtttaa	cacctgggtg	aggacctttg	ggtcctggcc	7020
atgtgtgtcc	ccttggtggt	tcccactgag	agctgttcgt	agataaatct	acataatacc	7080
aagtctggta	ttcatcatca	aaaacagcct	tccagccaga	aggcacttga	ggaggattac	7140
ttttactttg	agccattact	ttaaaatggt	ttgggttttt	ttatcgactc	tctgtagtgt	7200
agtataaaaa	gcaattcaga	aagagtcttt	tgattatcat	gactaaagta	aactttttct	7260
tccagctgcc	ggctcgggtc	ggtggacgaa	acaaaaaagc	ctgttgaaat	aatactgaat	7320
acgttaaaac	gtcacagcgc	ttgtatatct	tgtgagagaa	aggctattac	caactgttata	7380
tgaatataga	ttgtctcaaca	aagcgtgtct	ttttctgtgt	gcttaaaata	tagtgtcttt	7440
tttgggaaaa	aaacgtttat	ttacagtatt	ctcttctctc	tctctctatt	caactgtgtg	7500
cggttcgaag	aaaattacgc	ataaagaatc	aaagaaaggc	aaaaattacg	ctgtacgatg	7560
gggtcgtttc	ccttagctga	gtttccatta	cgtgatatcc	ctgttcctta	tagctaccgt	7620
gtgtctggcg	gtatagcttc	ctcaggtagt	gttactgcgc	ttgttactgc	cgctggcact	7680
catcgaaact	cgtccacggc	taagacagtt	gagacagaag	acggcgaaga	agatatcgat	7740
gagtatcaga	ggaaaagagc	agctggttct	ggcgaaatcca	ctcctgaacg	cagtgatctt	7800
aaaagggtaa	aacatgataa	tcacaaaacc	ctccatccag	ttaacttaca	gaacacgggt	7860
gcagcgtctg	tggataatga	cggctctgcac	aatttaacag	atataatcaa	cgatgcagaa	7920
aaacttttga	tgtctgtgga	tgatggttct	gccgcacctt	ctacattgag	tgtaaacatg	7980
ggagtggcat	ctcataatgt	tgctgctccc	actaccgtca	atgcggcaac	aataactggc	8040
agtgatgtta	gtaacaatgt	taatagtgtc	actattaaca	atcctatgga	ggaaggagcg	8100
ctgccgttat	caccactgc	ttcctctcca	ggtaccacaa	ctcctttagc	taaaactacg	8160
aaaactatca	acaacaataa	taatatcgcc	gatttgatag	aatccaaaga	ttctataatc	8220
tcccctgaat	acctttctga	tgagattttc	agcgcaataa	acaataatct	ccctcacgca	8280
tacttcaaaa	atttattatt	tagattagtt	gccaaatagg	ataggagtga	actatccgac	8340
ttggggactt	taatcaagga	taattttaag	agggacctaa	taacgtcttt	gccttttgaa	8400
ataagtttga	aaattttcaa	ttatttgcaa	ttcgaggata	ttataaattc	ccttgggggtc	8460
tcccaaaatt	ggaacaaaat	aattagaaaa	tctacatcgt	tgtggaaaaa	actttcgata	8520
tcggaaaatt	ttgtgagccc	aaagggtttt	aattctctca	atctcaaaact	ctcccaaaaa	8580
tacccaaaac	tctcacacaa	agatcgcttt	agattatctt	ttctggagaa	tatattcatt	8640
ttaaaaaatt	ggtacaatcc	caagtttgta	ccacaaagga	ccacgttaag	aggccatag	8700
acgagtgtta	ttacgtgctt	gcaatttgaa	gataattatg	tcattacggg	ggctgatgac	8760
aaaatgatca	gagtttatga	ttcgataaac	aagaaatttc	ttctacaact	atcagggtcat	8820
gatggtgggg	tttgggcatt	gaagtatgcc	catggcggta	ttttagttag	cggttctaca	8880
gacagaacgg	tgcgagtttg	ggatattaag	aaagggtgtt	gtacccatgt	gtttaaagggt	8940
cataactcta	cgggtgaggtg	cctagatata	gtagaatata	aaaatatcaa	gtacattgtt	9000
actgggtcga	gagataaac	tttgcacgtt	tggaatttgc	ccaaggagtc	ctccgttcct	9060
gatcatgggg	aagaacatga	ttatccatta	gtctttcata	cccctgagga	gaacccatat	9120
tttgttggtg	ttttaagagg	acatatggca	tctgtaagaa	ctgtctcagg	ccacggtaat	9180

attgtcgtta	gtggctccta	tgataataca	ctgattgtgt	gggatgttgc	gcaaatgaaa	9240
tgtttgtata	ttttaagtgg	acatacggat	cgtatttatt	cgacaatcta	cgatcatgaa	9300
agaaaaaggt	gcatctctgc	cagtatggat	accactatta	gaatttggga	tttggaataa	9360
atatggaata	atggagaatg	ttcctacgca	acaaattcag	catcgccatg	cgccaaaata	9420
cttggtgcta	tgtacacttt	gcagggtcat	acagctttgg	tcggtttatt	aagattatcc	9480
gacaaatttt	tggtcagtg	cgctgcagac	ggttcaataa	ggggttggga	cgcaaacgac	9540
tactctagaa	aattttccta	ccatcatacc	aatttgagt	caattaccac	attttatgta	9600
tcggataata	ttttggtgag	tggatcgga	aatcagttca	acatctataa	tctacggagt	9660
gggaaattgg	tccacgcaaa	tattctaaaa	gatgctgac	agatttgggc	ggttaatttt	9720
aagggcaaaa	cacttggtgc	agcagttgaa	aaagatggac	agagcttttt	agaaattctg	9780
gatttcagca	aagcttcaaa	aattaactac	gttagcaatc	ccgtaaaactc	ctcgtcgtcg	9840
tctttggaat	ccatttctac	ttctttgggt	ctaacgagga	caactataat	accatgacct	9900
ttcccagaga	ataagcattg	actcatactt	agataaata	gcttaataag	tagttatata	9960
atcagtaaaa	aagtacaata	acaacttcgt	acattttatt	gagtataaac	tgcagtaaaa	10020
ctgcctggat	gtgtcaattt	taattgtgtt	tacaaaaagg	gtgccgttta	tttaattaatg	10080
tttcttccct	gaaaatatgg	aaagtacaag	tttttagttg	agaagggttt	aagaaagttt	10140
tgaaaatgat	ctaaaaaaat	ataaaagcaa	tcaaagaaat	aaaagctgga	aaaatgcgta	10200
ataaccgaag	tgactaaaa	ttctttacgc	gccaaataag	aaatcgatgc	tcttgaaaagt	10260
agcaaccatt	ttttaataaa	tatatctctg	atggttcttg	gcccaagagt	tttcttgaac	10320
ttttaacggt	aagaagttga	ttctgctgat	tttttcaag	ttatcaagcg	ttatgttttg	10380
aaacatctgt	tccgttttca	gtttttcaaa	aagatcgata	ttctcttgag	aattaggaag	10440
ttctacgctc	aacctcccgt	gggcagaaga	aatacaaaag	ctaatacaat	tgtgttagaa	10500
taaagttcta	atattatcta	attagtagta	ttcatgttac	tagtatatta	tcacatgata	10560
tttctcaaat	tggacaatgt	aaataatgag	tgtttttatg	acacaatctt	tatcgttaga	10620
tgtttcaacg	ttccaaggct	tggttctatc	gctcttctct	tcaaattgta	atcgtttgtc	10680
atgataaaca	cgtacgggaa	aaaaaataaa	agtatcaata	acgcgtaaa	tgaatagagt	10740
attggattct	ataagaccga	agacgctcat	atcacatctc	ataaaatcac	ttaaagcaag	10800
catccagagg	ctattgataa	aaagcaggca	caaggagacg	caatgggacg	tttagttggc	10860
ttagaactaa	gtaatttcaa	gtcctataga	ggcggtacca	aggtaggatt	cggcgagtcg	10920
aatttcacaa	gtattatcgg	tcctaaccga	tctggtaaat	cgaatatgat	ggatgctatc	10980
tcattcgtac	tcgggtgctg	gagtaatcat	ttgaggtcaa	acatcttgaa	agattttatc	11040
tatagagggt	ttctaaccga	tgagaatagt	gacgattatg	ataacgaggg	cgctgcctct	11100
tcgaaccac	aatccgcata	cgtgaaggca	ttttatcaaa	aaggtaaaca	actggtggag	11160
ctgatgagga	taatttccag	aaacggtgac	actagttata	aaattgatgg	aaaaactgtc	11220
tcctataagg	actattctat	atttcttgag	aacgaaaata	ttcttatcaa	agccaaaaat	11280
ttctagtgt	tcagggtga	tgttgagcaa	attgcagcac	aatctcccgt	agaattatca	11340
agaatgtttg	aagaagtatc	aggttctatc	caatacaaaa	aggagtatga	agagttgaag	11400
gaaaagattg	agaaattaag	caaactctga	accgaatcta	taaaaaatag	gagaaggatc	11460
catggagaat	tgaagacata	taaagaaggt	atcaataaga	acgaggagta	taggaaacaa	11520
ttggacaaaa	aaaatgaatt	acagaagttc	caggctctat	ggcagttata	tcatttagag	11580
caacaaaaag	aggagctaac	ggacaagctg	tccgcattaa	actctgaaat	atcgtcttta	11640
aagggaaaaa	taaataacga	gatgaaatca	ttacaacgct	caaaatcttc	ctttgttaaa	11700
gaaagcgcag	taatttctaa	gcaaaaaagt	aaattagatt	atatcttcaa	ggataaggaa	11760
aaattagtct	cggatttacg	gctcataaag	gttcttcaac	aggcagcagg	gaagcgaatt	11820
tcccatatag	aaaaaagaat	cgaaagttta	cagaaagatc	ttcaaagaca	gaagacttat	11880
gtggagagat	ttgaaacaca	actaaaagt	gtgaccagat	caaagggaagc	ttttgaagag	11940
gaaatcaaac	aatctgctag	aaactatgac	aaattcaagc	taaatgaaaa	tgatttaaa	12000
acataataat	gcttacatga	aaaatatctg	actgaagggt	ggtcaatcct	agaagaaaaa	12060
attgccgttt	tgaacacga	taagcgagaa	atccaagagg	aattggagag	attcaacaaa	12120
agggcagata	tttctaaaag	aaggataacg	gaggagcttt	ctataacagg	agaaaaattg	12180
gacacgcaat	taaacgattt	aagagtttct	ttgaatgaga	aaaacgccct	tcatactgaa	12240
cgtttgcacg	agctgaaaaa	attacaatct	gatattgaat	ctgctaataa	tcaagaatac	12300
gacttaaat	tcaagttgag	ggaaacggtg	gttaagatcg	atgacttgag	tgctaataca	12360
agagaaacaa	tgaagaaaag	aaaactaaga	gaaaatatag	caatgttgaa	aagattcttc	12420
ccggagtaaa	aaggtcttgt	tcatgatctt	tgtcacccaa	aaaaggagaa	atatggcttg	12480
gcagtgtcta	ccatcttagg	taagaacttt	gattccgtca	ttgtagaaaa	tttaaccgta	12540
gctcaagaat	gcattgcatt	tttgaagaag	caacgtgcgg	gcactgcac	tttcatacca	12600
ctagacacaa	ttgagacaga	gttacctaca	ttatcattgc	ctgactcaca	agactatatt	12660
ttatcaatta	atgctattga	ctacgagccg	gaatatgaaa	aagcgatgca	atatgtgtgt	12720

ggcgattcca	tcatatgtaa	tacgttgaat	attgctaaa	atctgaaatg	gaaaaagggc	12780
ataagaggca	aattgggttac	aattgaaggt	gctttgatcc	acaagggcgg	tttcatgaca	12840
ggtggtatat	caggagatgc	caataatagg	tgggataaag	aagaatatca	aagcttaatg	12900
tctttaaaag	acaaattact	aatccaaatc	gatgaacttt	ccaacgggtca	acgctcta	12960
tcaatcagag	caagagaagt	tgaaaatagt	gtttcactat	tgaactcgga	catagcaa	13020
ttgagaactc	aagtaacaca	acaaaaacgc	tccttggtatg	aaaatcgttt	agagattaag	13080
taccataatg	acttgataga	gaaagaaatt	caaccgaaga	taactgaact	aaagaagaag	13140
ctagatgatt	tagaaaatac	taaagataat	ttagtgaag	agaaggaggc	tttacaaaat	13200
aatattttca	aagaattcac	tagtaaaatt	ggctttacaa	tcaaagaata	tgaaaatcat	13260
tccggtgaat	tgatgagaca	acaatctaaa	gaattacagc	agttacaaaa	acaaattttg	13320
accgttgaaa	ataagttgca	gtttgagaca	gacagactaa	gtactactca	aagaagatat	13380
gaaaaggcgc	aaaaggattt	agagaatgct	caagtgaaa	tgaagtcttt	ggaagaacag	13440
gaatatgcaa	tagaaatgaa	aatcggatca	atagagtcta	aattggaaga	acacaaaaat	13500
cacttagatg	agttacagaa	gaaatttgta	acgaagcaaa	gtgaattaaa	ttccagcgaa	13560
gatattctag	aggacatgaa	cagcaactta	caagtcctaa	aaagggaag	agacggtata	13620
aaggaagata	ttgaaaagtt	tgatttagag	agagtaacag	cgttaaagaa	ttgtaaaatt	13680
tctaataata	atatacctat	atcatctgaa	acaacgatag	atgattttacc	aatatcttcc	13740
actgataatg	aagcaattac	aatttccaac	agtatcgata	taaactataa	aggactacct	13800
aaaaaataca	aagaaaacaa	taccgattcg	gcaaggaagg	agctggagca	gaagattcat	13860
gaagtggagg	aaatatggaa	cgagttgcag	cccaatgcaa	gagctttgga	gagatacgac	13920
gaggcggaag	gaagggttga	agtgattta	aacgaacacag	aacaactaaa	ggccggaagaa	13980
aagaaaatat	taaaccagtt	cctaaaaaatt	aagaaaaaaa	gaaaggaact	gttcgaaaag	14040
acatttgatt	atgtgagcga	ccatttagac	gcaatctaca	gggaacttac	taaaaaatccc	14100
aactccaatg	tggaattggc	cggtggtaac	gcttctttaa	ccatagaaga	cgaagatgaa	14160
ccgttcaatg	cggaatcaa	atatcatgcc	actccgcctc	ttaaaagatt	caaagacatg	14220
gaatatcttt	ctggtgggtga	aaaaaccgta	ctgcatattg	gtctattatt	tgctattaat	14280
tcctaccagc	ctagtccttt	cttcgtgctg	gacgaagtgg	acgcagccct	agacattact	14340
aacgtccaga	gaattgctgc	ctatataaga	aggcaccgta	atccagatct	ccaattcatt	14400
gtcatttcat	tgaagaacac	catgtttgaa	aaatctgacg	ctctcgttgg	ggtttacaga	14460
cagcaacaag	aaaactcgtc	gaagatcata	actttggact	tgagcaatta	cgcagaataa	14520
tctatgaaac	caacctctgc	tataaccctg	caaataacta	ataatatcta	tatagggtcaa	14580
ctagctagt	caatatcata	gtaacaataa	tattaaatac	gtcacttttt	ttccagggtta	14640
acccaattgt	ggtgggtggc	ggccgaggtta	tccttagaa	aagaattttt	taagttcttt	14700
ctcatctctt	accagtgagg	aagtacacga	tatttgcaaa	gtctgtcatc	agggcttgat	14760
aataaagctg	cattagatct	tagcaaaaac	tacgagaaga	acattgaata	ttgtagctgt	14820
atttgcatac	ataaaacttta	tcattgttcg	ttagctagct	ttgcacatta	atttttcgat	14880
ttgttaccgc	caatgaccgc	taacaatgac	gatgatatac	aatcacccat	tcccattact	14940
aacaagacct	tactcccaatt	gaagcgcttt	gagagaagtc	caggaaggcc	cagttcttct	15000
cagggcgaga	taaaacgtaa	aaagtctagg	ctatatgccg	cagacggaag	accacattct	15060
ccgctaagag	caaggtctgc	taccccaacg	ctacaggacc	aaaaactgtt	caatggcatg	15120
gattccaactt	cccttttgaa	tgaaaggcta	cagcattata	cgctggatta	tgttagtgc	15180
agggcccagc	atatgaagaa	tatatatgac	ccatcatctc	gctggttcag	cagatccgtg	15240
aggcctgaat	ttcctattga	agagttctta	ccatataaga	ctgaaagtca	tgaagatcaa	15300
gcgaaatact	tgtgccatgt	cttagttaat	cttttatattg	cgatcagctc	attagatata	15360
caaggtttga	tttctatttc	cagtaaagat	ctggctgatt	taaagaaaga	agtggtgat	15420
ttagctctta	aaaccgatct	tttcagggtta	tctaacaaca	cagcggagaa	tgacttactt	15480
ggtaacgata	ttgctgatta	tgacgatcgc	gaaggcctgg	aggacgaatt	ggatgaatac	15540
ttcgacttag	caggccccga	ctttaatgct	accggaaaaa	tcacagctaa	atcagctact	15600
attgtgaatg	taaaccattg	gactaatgag	ctcaaaaatt	gtctacattt	tgattttcca	15660
gtagctttta	gaaagtcgct	agcgacaggt	tattattatt	tgtctcttgt	tcagggccaa	15720
aaggtgtata	gacaaatgca	tgtcgatatg	ttcgaaagat	tagtaagcct	tgacgatgat	15780
aggacaaatt	tcactgaact	gttgcaaaaa	caaggtcttt	tgctagatca	tcaaatcatg	15840
ctcaatttcc	tgtgcgaatt	tctaccttat	ccagatcccc	actatgctcg	ttatgaatta	15900
tcatcaaagg	ataagtttaca	gttattttaga	ttacttttga	agcatgcaca	taatgcaaaa	15960
ccatttttgc	ataagtcaaa	ggaaagtgtg	ttagttgata	cgatgaattt	tctgttgtct	16020
agtcttgca	catctactat	gatggctgtc	atgcctatcg	ttacatccgt	tgtgccctat	16080
cattatcata	tccattctaa	gatcatcgat	tatttcccgt	tttgctatag	catctggagc	16140
tcagtgcgcg	caaacgtggc	catcgacact	cacatgtatg	atgttgttgg	gtcaatttcc	16200
aaagacgttc	ataataagat	tttgagtgc	gagcatgaaa	aggatgtagt	tggagtggag	16260
tttgagagaat	tcgggatttt	tactgatgat	caaatagact	ttatgttcaa	taggttacia	16320

ggccatctta	gaacagacgg	tcaaatacat	tcgatttccc	gcacagtgaa	gccttttgtt	16380
tatgctataa	acggatcaaa	aaaagatagg	ttttttgaaa	aacttgtaag	tttagccaaa	16440
gcaatcgaaa	catttatcca	tccctctaatt	aatgggtttt	ggaccaagcc	aaatgctaaa	16500
ttcgttcatg	catttataaa	gtcttaccac	ggaaggggtca	aatatgaaga	agatatttgt	16560
gctagggggc	tcacaaatgg	gatatgttta	acttcttttt	gtcacgaaga	gatagttgaa	16620
atattcttaa	atattatcag	tctgggttca	cagaataaaa	atcctgatat	tgcgaactat	16680
tacatctctt	gttttcgcata	tctgttagag	ctggatcctt	caaatgcata	tttaattttat	16740
gacaaaatac	tgatagattt	gtacgatacg	ctggctgacc	aatttatcaa	ttcgagacac	16800
agaatcattt	cctctttgaa	acaattttaca	agagtaattc	ggttttattgt	gatggataag	16860
ctatatcgcg	tgcacataac	aaacgtcctt	tcgatgctgg	tctccaaact	tgatatgaac	16920
gatactaatt	tgacaagcaa	cctcatcaac	ggtattgtat	ctatagccgc	tttcattcct	16980
atccaagacc	tactggggga	ggacgattat	atcgtgttg	aatcggatac	tcttctttta	17040
gttcaacaac	atttttatca	tattaagtgt	ggcgaaagtt	caaagacctt	ccgagttgat	17100
gatgaactgt	taaataacgc	tttcaaagct	tccactacag	tttttcaaag	tatgctaaaa	17160
gtatacgtag	aaaaaatttt	ccaattgggt	gatgtagact	tagaggactc	tttggttact	17220
aaaataaacc	aaacaactat	gatttttaca	gagtctatgg	acgataaaat	attttaattat	17280
tttgcttctt	tattacagag	aaacttctgg	agtaatgact	ccttcaagga	aaaggatcca	17340
aactatgaat	tagttactat	cccactagcg	gcttttagtaa	gaaggataaa	tgggttgagt	17400
aaagagttgg	tcagaactct	tttattccat	atcaaagaac	aatcaaaaag	aggcgccggg	17460
tctgtaagaa	gtacttcaga	gattcaacag	agggatgtta	agttagtttt	atacttaact	17520
gcactaaatg	atgtcttaag	gcaatgtcat	gaatctctat	tggagtatag	cgatgagttg	17580
ataacattca	tgaaatattt	atacgacaac	gtcactaacc	cgccattgga	tgttatcaca	17640
tctattgtta	ttcacagtgc	tttagcaact	ctatgtacaa	ccgaaataac	tgattgtcgt	17700
ctattcccag	aggactctaa	gattccggaa	aaagacagat	ggggaggact	acagttcgat	17760
cctcgaagat	ttgataaaca	gcatttgagt	tttcagtggc	acgtaccttc	tagtgacgag	17820
ataactttat	ccataagcat	tctagaaagc	ctttccgaat	actgtattaa	caacgtagaa	17880
gaactgatga	aagctccaag	acatgattcc	gaatatggcg	atatgataca	aaaaatagtt	17940
ttagtattga	cacatacgct	ttccgggtca	agtttacttt	ttgatccaga	ttttaacaaa	18000
tataggacgc	aatcaaactt	atcatacaga	gagaaactga	ttttattaaa	gaatatacgc	18060
gaaaataaatt	gtgaccctca	agaactggat	attgatattg	aacaaattcg	ttctggcaag	18120
gatgatgaag	catatattga	gagtaaggac	attgaagcag	ggctgaatgc	aggagtttcc	18180
gatgttgtgc	agtttaagaga	tgagtttccg	gatgaattaa	ttggtgatga	agaagtagtg	18240
tctgagatgc	catctgggtg	aaataccctt	atcgcgggga	cgcatggcac	ggacaattca	18300
gctatgagtt	cggatctagc	tttcagggat	ttagatattt	acacctgtaa	ttattacttc	18360
ggaaataacca	ctgaggagaa	gttacaaaac	ccacaatatt	tacaagtcca	cagagttaga	18420
gcgcgcattg	gacattttct	tcacaaactc	tatgtttttt	tatctacaaa	cttcgaaaaac	18480
aacaccaaca	tgttccaaat	tctattgcac	ggattgaaag	tttggttcac	agatctggga	18540
caagaaaacgg	tcttcaatga	agacccaaat	gccttcattg	acgttgattt	cctagaaaaac	18600
gttcaatctc	tctcacacgt	aaatgagccc	ttcacgagaa	ccaattttgc	aatcagagca	18660
aacagtttgc	accagagtag	agttctatta	cattcaacaa	atagaaaagc	ttccaagctg	18720
gaaaacctat	tgttagttga	catcatacag	ttagcgacat	ccctttatcc	tgatattttat	18780
aaaccagcac	agggaaactt	ggtacactgt	atgaaacaat	tagttgggtc	atatggcgta	18840
gtcatcaata	aaattattcc	atcatttagag	aaagcgatta	aggatcatga	ttatatgaaa	18900
atccaagtta	ttttaaatgt	tttgtttaatt	aagaaaatcc	ataggaagct	tatgacggat	18960
tataaagaca	tcggcagatt	gatatttctg	cttattgaaat	gttgctcgtg	gaatgaatta	19020
gagattggta	tgtatgcaga	taaaatctta	actgacatag	tgattgggat	caagattcct	19080
tctagtgtat	gtgtcatttc	cgatcaagct	ttcttacctt	tagcacctcc	tgatgggtact	19140
attaatttgc	aagttgaagc	ggtaaaagct	gccaaaaaga	aaaagcgtga	gtactacctc	19200
tctctgttag	tggtattgca	ggacaaaact	ttagacaaat	tagataatga	aaaagataatg	19260
gggtggaaga	taagaatggt	cattttacgt	tttgttacac	aaatccaatc	gaacctcgaa	19320
agcaaaccgc	ataaaagagc	agtattttca	ataatctccc	aaatctccac	aaaacatccc	19380
gaaatcatac	atgttggtgt	aaagtcatgt	ttgtcgacgt	gcaacaagat	aatatctctc	19440
tctgactatg	aatatgacat	caccagggcc	tataagaatg	aattcaatcc	atcattttgtg	19500
gaaatactgg	acacttcgac	cacaagcttc	cctaaaaacgt	ttactgaaga	aatgaataac	19560
tttgataaac	ccaagtattt	tatcgatttg	agggcgtagt	taggggtggct	atgttgggga	19620
aggcttatgt	acgtcatgtc	gccgaaagct	ttaaagctca	atttacgtga	gaatgaactg	19680
gaagtcctca	agacagctgg	tcatctattg	acaagagaat	tcctgagaga	tgttacaatg	19740
aatttagtcc	aagataatga	aactaggggt	gttttttagta	gtggttaacgt	gtcattttttc	19800
tcttttagtaa	tccttttgat	atcatctggg	ttctgcgaac	tgaatatgtc	ggatctctttt	19860

```

gagctatgtg aatcctacta taacaaagac gataaggott cgatgatcat gtctgtcgag 19920
atagtggctg gcttagtttg cgggagtaag tttatgtcag tctctgactt ggacaaacgt 19980
gacactttta tcgaaaactt cctagccaaa tgttttagatt atgagttgaa ccatgacgca 20040
tttgaaattt ggagcacctt ggcatgggtg ttgcctgcag tcggtgattt aagaagggtct 20100
aaaacttttt ttgcccattt tatcaacgcc gatggcatgt ttgaccgtga atctgatgca 20160
gccacacatc aaacctccaa aattttacatg ctaagaagta tcttgatgag catggaattt 20220
agagccccag atgttggtaa gctatttgat gagttgggtat ttgatcacc atacgatcag 20280
gttcgccagc tgtcgctaaa ctattgacga ccttagttca aaatcaaagt aatccgtcaa 20340
tttcagatcc aaccacatta ttagaagcag aacggaatga tcc 20383

```

<210> 2

<211> 779

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 2

```

Met Gly Ser Phe Pro Leu Ala Glu Phe Pro Leu Arg Asp Ile Pro Val
  1                      5                      10                      15

Pro Tyr Ser Tyr Arg Val Ser Gly Gly Ile Ala Ser Ser Gly Ser Val
                      20                      25                      30

Thr Ala Leu Val Thr Ala Ala Gly Thr His Arg Asn Ser Ser Thr Ala
  35                      40                      45

Lys Thr Val Glu Thr Glu Asp Gly Glu Glu Asp Ile Asp Glu Tyr Gln
  50                      55                      60

Arg Lys Arg Ala Ala Gly Ser Gly Glu Ser Thr Pro Glu Arg Ser Asp
  65                      70                      75                      80

Phe Lys Arg Val Lys His Asp Asn His Lys Thr Leu His Pro Val Asn
                      85                      90                      95

Leu Gln Asn Thr Gly Ala Ala Ser Val Asp Asn Asp Gly Leu His Asn
  100                      105                      110

Leu Thr Asp Ile Ser Asn Asp Ala Glu Lys Leu Leu Met Ser Val Asp
  115                      120                      125

Asp Gly Ser Ala Ala Pro Ser Thr Leu Ser Val Asn Met Gly Val Ala
  130                      135                      140

Ser His Asn Val Ala Ala Pro Thr Thr Val Asn Ala Ala Thr Ile Thr
  145                      150                      155                      160

Gly Ser Asp Val Ser Asn Asn Val Asn Ser Ala Thr Ile Asn Asn Pro
                      165                      170                      175

Met Glu Glu Gly Ala Leu Pro Leu Ser Pro Thr Ala Ser Ser Pro Gly
                      180                      185                      190

Thr Thr Thr Pro Leu Ala Lys Thr Thr Lys Thr Ile Asn Asn Asn Asn
  195                      200                      205

Asn Ile Ala Asp Leu Ile Glu Ser Lys Asp Ser Ile Ile Ser Pro Glu
  210                      215                      220

```

Tyr Leu Ser Asp Glu Ile Phe Ser Ala Ile Asn Asn Asn Leu Pro His
 225 230 235 240
 Ala Tyr Phe Lys Asn Leu Leu Phe Arg Leu Val Ala Asn Met Asp Arg
 245 250 255
 Ser Glu Leu Ser Asp Leu Gly Thr Leu Ile Lys Asp Asn Leu Lys Arg
 260 265 270
 Asp Leu Ile Thr Ser Leu Pro Phe Glu Ile Ser Leu Lys Ile Phe Asn
 275 280 285
 Tyr Leu Gln Phe Glu Asp Ile Ile Asn Ser Leu Gly Val Ser Gln Asn
 290 295 300
 Trp Asn Lys Ile Ile Arg Lys Ser Thr Ser Leu Trp Lys Lys Leu Leu
 305 310 315 320
 Ile Ser Glu Asn Phe Val Ser Pro Lys Gly Phe Asn Ser Leu Asn Leu
 325 330 335
 Lys Leu Ser Gln Lys Tyr Pro Lys Leu Ser Gln Gln Asp Arg Leu Arg
 340 345 350
 Leu Ser Phe Leu Glu Asn Ile Phe Ile Leu Lys Asn Trp Tyr Asn Pro
 355 360 365
 Lys Phe Val Pro Gln Arg Thr Thr Leu Arg Gly His Met Thr Ser Val
 370 375 380
 Ile Thr Cys Leu Gln Phe Glu Asp Asn Tyr Val Ile Thr Gly Ala Asp
 385 390 395 400
 Asp Lys Met Ile Arg Val Tyr Asp Ser Ile Asn Lys Lys Phe Leu Leu
 405 410 415
 Gln Leu Ser Gly His Asp Gly Gly Val Trp Ala Leu Lys Tyr Ala His
 420 425 430
 Gly Gly Ile Leu Val Ser Gly Ser Thr Asp Arg Thr Val Arg Val Trp
 435 440 445
 Asp Ile Lys Lys Gly Cys Cys Thr His Val Phe Lys Gly His Asn Ser
 450 455 460
 Thr Val Arg Cys Leu Asp Ile Val Glu Tyr Lys Asn Ile Lys Tyr Ile
 465 470 475 480
 Val Thr Gly Ser Arg Asp Asn Thr Leu His Val Trp Lys Leu Pro Lys
 485 490 495
 Glu Ser Ser Val Pro Asp His Gly Glu Glu His Asp Tyr Pro Leu Val
 500 505 510
 Phe His Thr Pro Glu Glu Asn Pro Tyr Phe Val Gly Val Leu Arg Gly
 515 520 525
 His Met Ala Ser Val Arg Thr Val Ser Gly His Gly Asn Ile Val Val
 530 535 540

Ser Gly Ser Tyr Asp Asn Thr Leu Ile Val Trp Asp Val Ala Gln Met
 545 550 555 560
 Lys Cys Leu Tyr Ile Leu Ser Gly His Thr Asp Arg Ile Tyr Ser Thr
 565 570 575
 Ile Tyr Asp His Glu Arg Lys Arg Cys Ile Ser Ala Ser Met Asp Thr
 580 585 590
 Thr Ile Arg Ile Trp Asp Leu Glu Asn Ile Trp Asn Asn Gly Glu Cys
 595 600 605
 Ser Tyr Ala Thr Asn Ser Ala Ser Pro Cys Ala Lys Ile Leu Gly Ala
 610 615 620
 Met Tyr Thr Leu Gln Gly His Thr Ala Leu Val Gly Leu Leu Arg Leu
 625 630 635 640
 Ser Asp Lys Phe Leu Val Ser Ala Ala Ala Asp Gly Ser Ile Arg Gly
 645 650 655
 Trp Asp Ala Asn Asp Tyr Ser Arg Lys Phe Ser Tyr His His Thr Asn
 660 665 670
 Leu Ser Ala Ile Thr Thr Phe Tyr Val Ser Asp Asn Ile Leu Val Ser
 675 680 685
 Gly Ser Glu Asn Gln Phe Asn Ile Tyr Asn Leu Arg Ser Gly Lys Leu
 690 695 700
 Val His Ala Asn Ile Leu Lys Asp Ala Asp Gln Ile Trp Ser Val Asn
 705 710 715 720
 Phe Lys Gly Lys Thr Leu Val Ala Ala Val Glu Lys Asp Gly Gln Ser
 725 730 735
 Phe Leu Glu Ile Leu Asp Phe Ser Lys Ala Ser Lys Ile Asn Tyr Val
 740 745 750
 Ser Asn Pro Val Asn Ser Ser Ser Ser Ser Leu Glu Ser Ile Ser Thr
 755 760 765
 Ser Leu Gly Leu Thr Arg Thr Thr Ile Ile Pro
 770 775

<210> 3

<211> 2151

<212> DNA

<213> Homo sapiens

<400> 3

tgcgttggt ggcgcctggc accaaagggg cggccccggc ggagagcgga cccagtggcc 60
 tcggcgatta tggaccggc cgaggcgtg ctgcaagaga aggcactcaa gtttatgaat 120
 tcctcagaga gagaagactg taataatggc gaaccccta ggaagataat accagagaag 180
 aattcactta gacagacata caacagctgt gccagactct gcttaaacca agaaacagta 240
 tgtttagcaa gcaactgctat gaagactgag aattgtgtgg ccaaaacaaa acttgccaat 300
 ggcacttcca gtatgattgt gcccaagcaa cggaaactct cagcaagcta tgaaaaggaa 360
 aaggaactgt gtgtcaata ctttgagcag tggtcagagt cagatcaagt ggaatttgtg 420

```

gaacatctta tatcccaaat gtgtcattac caacatgggc acataaaactc gtatcttaaa 480
cctatgttgc agagagattt cataactgct ctgccagctc ggggattgga tcatatcgct 540
gagaacattc tgtcatacct ggatgccaac tcaactatgtg ctgctgaact tgtgtgcaag 600
gaatggtacc gagtgacctc tgatggcatg ctgtggaaga agcttatcga gagaatgggc 660
aggacagatt ctctgtggag aggcctggca gaacgaagag gatggggaca gtattttattc 720
aaaaacaaac ctctgacgg gaatgctcct cccaactctt tttatagagc actttatcct 780
aaaattatac aagacattga gacaatagaa tctaattgga gatgtggaag acatagttta 840
cagagaattc actgccgaag tgaaacaagc aaaggagttt actgtttaca gtatgatgat 900
cagaaaatag taagcggcct tcgagacaac acaatcaaga tctgggataa aaacacattg 960
gaatgcaagc gaattctcac aggccatata gggttcagtc tctgtctcca gtatgatgag 1020
agagtgatca taacaggatc atcggattcc acggtcagag tgtgggatgt aaatacaggt 1080
gaaatgctaa acacgttgat tcaccattgt gaagcagttc tgcacttgcg tttcaataat 1140
ggcatgatgg tgacctgctc caaagatcgt tccattgctg tatgggatat ggcctcccca 1200
actgacatta cctcctggag ggtgctggtc ggacaccgag ctgctgtcaa tgtttagtag 1260
tttgatgaca agtacattgt ttctgcatct ggggatagaa ctataaaggt atggaacaca 1320
agtacttggt aatttgtaag gaccttaaat ggacacaaac gaggcattgc ctgtttgcag 1380
tacagggaca ggctggtagt gagtggctca tctgacaaca ctatcagatt atgggacata 1440
gaatgtggtg catgtttacg agtgtagaa ggccatgagg aattggtgcg ttgtattcga 1500
tttgataaca agaggatagt cagtggggcc tatgatggaa aaattaaagt gtgggatctt 1560
gtggctgctt tggacccccg tgctcctgca gggacactct gtctacggac ccttgtggag 1620
cattccggaa gagtttttcg actacagttt gatgaattcc agattgtcag tagttcacat 1680
gatgacacaa tcctcatctg ggacttcta aatgatccag ctgcccgaagc tgaaccccc 1740
cgttccccct ctcgaaacata cacctacatc tccagataaa taaccataca ctgacctcat 1800
acttgcccag gaccatttaa agttgcggtt ttaacgtat ctgccaatac caggatgagc 1860
aacaacagta acaatcaaac tactgccagc tttccctgga ctagccgagg agcagggtct 1920
tgagactcct gttgggacac agttggtctg cagtgggccc aggacggctc actcagcaca 1980
actgactgct tcagtgctgc tatcagaaga tgtcttctat caattgtgaa tgattggaac 2040
ttttaaacct cccctcctct cctcctttca cctctgcacc tagttttttc ccattgggtc 2100
cagacaaagg tgacttataa atatatttag tgttttgcca gaaaaaaaaa a 2151

```

<210> 4

<211> 569

<212> PRT

<213> Homo sapiens

<400> 4

Met Asp Pro Ala Glu Ala Val Leu Gln Glu Lys Ala Leu Lys Phe Met
1 5 10 15

Asn Ser Ser Glu Arg Glu Asp Cys Asn Asn Gly Glu Pro Pro Arg Lys
20 25 30

Ile Ile Pro Glu Lys Asn Ser Leu Arg Gln Thr Tyr Asn Ser Cys Ala
35 40 45

Arg Leu Cys Leu Asn Gln Glu Thr Val Cys Leu Ala Ser Thr Ala Met
50 55 60

Lys Thr Glu Asn Cys Val Ala Lys Thr Lys Leu Ala Asn Gly Thr Ser
65 70 75 80

Ser Met Ile Val Pro Lys Gln Arg Lys Leu Ser Ala Ser Tyr Glu Lys
85 90 95

Glu Lys Glu Leu Cys Val Lys Tyr Phe Glu Gln Trp Ser Glu Ser Asp
100 105 110

Gln Val Glu Phe Val Glu His Leu Ile Ser Gln Met Cys His Tyr Gln
115 120 125

His Gly His Ile Asn Ser Tyr Leu Lys Pro Met Leu Gln Arg Asp Phe
 130 135 140
 Ile Thr Ala Leu Pro Ala Arg Gly Leu Asp His Ile Ala Glu Asn Ile
 145 150 155 160
 Leu Ser Tyr Leu Asp Ala Lys Ser Leu Cys Ala Ala Glu Leu Val Cys
 165 170 175
 Lys Glu Trp Tyr Arg Val Thr Ser Asp Gly Met Leu Trp Lys Lys Leu
 180 185 190
 Ile Glu Arg Met Val Arg Thr Asp Ser Leu Trp Arg Gly Leu Ala Glu
 195 200 205
 Arg Arg Gly Trp Gly Gln Tyr Leu Phe Lys Asn Lys Pro Pro Asp Gly
 210 215 220
 Asn Ala Pro Pro Asn Ser Phe Tyr Arg Ala Leu Tyr Pro Lys Ile Ile
 225 230 235 240
 Gln Asp Ile Glu Thr Ile Glu Ser Asn Trp Arg Cys Gly Arg His Ser
 245 250 255
 Leu Gln Arg Ile His Cys Arg Ser Glu Thr Ser Lys Gly Val Tyr Cys
 260 265 270
 Leu Gln Tyr Asp Asp Gln Lys Ile Val Ser Gly Leu Arg Asp Asn Thr
 275 280 285
 Ile Lys Ile Trp Asp Lys Asn Thr Leu Glu Cys Lys Arg Ile Leu Thr
 290 295 300
 Gly His Thr Gly Ser Val Leu Cys Leu Gln Tyr Asp Glu Arg Val Ile
 305 310 315 320
 Ile Thr Gly Ser Ser Asp Ser Thr Val Arg Val Trp Asp Val Asn Thr
 325 330 335
 Gly Glu Met Leu Asn Thr Leu Ile His His Cys Glu Ala Val Leu His
 340 345 350
 Leu Arg Phe Asn Asn Gly Met Met Val Thr Cys Ser Lys Asp Arg Ser
 355 360 365
 Ile Ala Val Trp Asp Met Ala Ser Pro Thr Asp Ile Thr Leu Arg Arg
 370 375 380
 Val Leu Val Gly His Arg Ala Ala Val Asn Val Val Asp Phe Asp Asp
 385 390 395 400
 Lys Tyr Ile Val Ser Ala Ser Gly Asp Arg Thr Ile Lys Val Trp Asn
 405 410 415
 Thr Ser Thr Cys Glu Phe Val Arg Thr Leu Asn Gly His Lys Arg Gly
 420 425 430

Ile Ala Cys Leu Gln Tyr Arg Asp Arg Leu Val Val Ser Gly Ser Ser
 435 440 445
 Asp Asn Thr Ile Arg Leu Trp Asp Ile Glu Cys Gly Ala Cys Leu Arg
 450 455 460
 Val Leu Glu Gly His Glu Glu Leu Val Arg Cys Ile Arg Phe Asp Asn
 465 470 475 480
 Lys Arg Ile Val Ser Gly Ala Tyr Asp Gly Lys Ile Lys Val Trp Asp
 485 490 495
 Leu Val Ala Ala Leu Asp Pro Arg Ala Pro Ala Gly Thr Leu Cys Leu
 500 505 510
 Arg Thr Leu Val Glu His Ser Gly Arg Val Phe Arg Leu Gln Phe Asp
 515 520 525
 Glu Phe Gln Ile Val Ser Ser Ser His Asp Asp Thr Ile Leu Ile Trp
 530 535 540
 Asp Phe Leu Asn Asp Pro Ala Ala Gln Ala Glu Pro Pro Arg Ser Pro
 545 550 555 560
 Ser Arg Thr Tyr Thr Tyr Ile Ser Arg
 565

<210> 5

<211> 4470

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 5

```

cacgctaatac atgactcaaa taaatccata aagttttata cagtttttaa aaatatcatc 60
atctattacc cagatgtgtt aatgaaccat tctatagtca tcattactag gctttcatta 120
ctactgaggt taccgccta tggaccttcc ttggtaaagg aacttgtttt aaaatttgcc 180
ttttaacaaa tttagtata tgattatcaa aaaaggcgtg gcaaaatata taacaccaa 240
tttaactgtg cctgtgtgtt actttctttt gtccatactt caccagtttt tcgattttac 300
acaataattc gttttcattt aatcgttctc ttagaagccc ggtttttgaa tatcaaaatc 360
gtacttggtg ccaactagca gggaagccca aaaattaagg cattgcattt aagcttacac 420
ctgcgtgaaa tcttgaaatt tctcattgat ttccggcaca taattatcat tggtagtgag 480
gctaaacagt ttgcggtttc ctttatacta agaaggctca taatggatca ggataacaac 540
aaccacaatg acagcaatag gctgcaccca cctgatatac atccaaattt gggccctcaa 600
ttgtggctga atagtagcgg tgattttgac gacaacaaca acaacaaca caacaacaac 660
aataataata gcacaagacc acaaatgcc tcaagaaact gagaaacggc aacttcggaa 720
agaaatgcaa gtgaggttag ggatgcaacg ctaaataata tctttagggt cgatagtatt 780
caacgggaaa cgcttttgcc aaccaacaac ggacaaccgc taaatcaaaa cttttcgctg 840
acatttcaac cacaacagca aacaaatgcg ctgaacggga ttgacataaa cactgtgaac 900
acaaacctta tgaatgggtg caatgttcaa atagatcaac ttaatcgatt gttaccgaac 960
ctaccagagg aagaacggaa gcaaatccac gaattcaagc taatagtggg caaaaaaatc 1020
caagagtttc tgggtgttat agagaaacgt agaaaaaaa tactgaacga aattgagcta 1080
gacaacctta aactaaagga gctacgtatt gataactccc cacaagcaat tagttatttg 1140
cataaattac aaagaatgag gcttagggcg ctagagacag aaaacatgga aattagaaat 1200
ttaaggctaa aaatattaac aattatagaa gagtacaaaa agtcattata tgcatactgc 1260
cattccaagc taagaggtca acaagtggaa aatccaacag ataatttcac catttgata 1320
aactccatag atactactga atcatctgac ttgaaagaag ggctacaaga tctttcgaga 1380
tattcaaggc agttcataaa taatgtgctt tcgaatccat caaatcaaaa catatgtacg 1440
agtgtcacc gaagatcacc tgtgtttgcc ctaaacatgc taccctcgga aatattacac 1500
ttaatattag ataaacttaa ccaaaaatat gatattgtaa aattccttac cgtttccaaa 1560

```

```

ctctgggctg aaataattgt gaagatactt tattacagac cgcacatcaa caaaaagagt 1620
caattagact tgtttttaag gactatgaag ttaacttctg aagaaactgt attcaactat 1680
cgtttaatga tcaaaagatt aaatthttca tttgttggtg actacatgca cgatacagag 1740
cttaactatt ttgtcggatg taagaatttg gagcgactaa ctttagtatt ttgcaagcat 1800
ataaccagtg ttccaatatc ggctgttttg agaggatgta aattttctcca aagtgtggat 1860
atcactggaa taagagacgt ttccgatgac gtatttgata ccttagcgac atattgtccc 1920
agagtacagg gcttttatgt tcctcaggca aggaatgtaa cattcgattc actgcggaat 1980
ttcatagtcc attccccgat gttgaaaaga ataaaaatca cagcaaacaa taacatgaat 2040
gacgaattag tagaactatt agccaacaaa tgccctttgc ttgtagaggt cgatataaca 2100
ttaagtccaa atgtcactga ttctagtttg ttaaaactcc tcactagggt agttcagctg 2160
agggaattca gaataactca taatacgaat attacggata atcttttcca ggagctttct 2220
aaagtagttg acgatatgcc ctctttaaga ttgattgatc tttctggatg tgaaaatatt 2280
acagataaaa ctatagaaaag tatagtcaat ttagccctta aattacgtaa tgtttttcta 2340
ggcaagtgtg gccgaattac agatgcatcg ttgttccaat tatcgaagct gggcaaaaac 2400
ttgcaaacag tgcatttttg gcactgtttc aatataactg ataacggggt aagagcactc 2460
tttcattcat gtacaagaat acagtatgtg gactttgcgt gctgtacgaa ttttaaccaat 2520
agaactcttt atgaactagc agacttacca aaattaaaga gaattggcct tgtcaaatgt 2580
acgcaaatga ctgacgaggg tttgttgaat atgggttccct tgcgaggccg aaatgatact 2640
ttagaagggg tacatthtatc ttactgttct aatttaacaa tatatccgat atatgagctt 2700
ctaattgtctt gcccaaggct ctcacatttg tctttgactg ctgttccgtc atttttacgc 2760
cccgatataa cgatgtattg caggcctgca ccctcagact ttagtgaaaa tcaacgtcaa 2820
atattctgcg tattttcagg gaaagggttt cataaacttc gccattattht agtaaattht 2880
acgtcgccccg cttttggacc acatgtcgat gtaaatgatg ttttgacaaa atatattaga 2940
tccaagaatt tgatattthaa cgggtgaaaca cttgaagatg ctcttaggag aatcataact 3000
gattthaaatc aagattccgc tgcaattata gctgctacag gattaaatca aatcaacggt 3060
ctaaataacg atthttcttht ccagaatatc aatthtgaac gaatagatga agtattcagt 3120
tggatctca atactthtga tggcattagg atgagctcgg aggaagttaa ctactatta 3180
ttgcaagtaa acaagacgtt ttgtgaagat ccatttagtg atgtggacga tgatcaagat 3240
tatgtcgtag cacctgggtg aaaccgggaa attaacagtg aaatgtgtca tattgttaga 3300
aaattccatg agttaaatga tcatattgat gatttcgagg tgaatgttgc tagtttggtg 3360
agagttcagt ttcagthttac tggthtttht ctctcatgaaa tgactcaaac ctatattgcaa 3420
atgattgaat taaacagaca aatthgttht gtacaaaaaa cggttcagga atcgggcaac 3480
atagattacc aaaaagggtc thtaatatgg cgactthttat tcattgacaa attcattatg 3540
gtggttcaga agtacaagct ctccaccgtt gthttgagac tatattthaaa agataacata 3600
acattgttaa ccagacaaaag agaactatta atagcccacc aaagatcagc atggaataac 3660
aataatgaca atgacgccaa ccggaacgcc aacaacatag tgaatattgt atcggatgct 3720
ggggcaaacg atacaagtaa caatgaaact aacaatggta atgatgacaa tgaaacagaa 3780
aatccaaatt tctggcgctc gthttggcaat agaattgcaa tatcacctga ccagatgagg 3840
aatctccaaa tgggacttcg taatcagaac atggttagga acaataacaa caacacaatt 3900
gacgaatcaa tgcttgacac tgccattgat tctcaaatgg atgaagcatc aggaacgccc 3960
gatgaagata tgttataatt gtatttcatt gaatacttac tgtcctacta cacctthtatt 4020
ttcaaaatcc cactthttctt acttatttac atataaatac ataatgcatt cactthtgaaa 4080
ctthttgcct tagcatgtat acgctataga cttgcggtat caacgaatat acgtaacgtt 4140
gtcacgtcca cagaagatgc tatgtcaaca gttccctgca gatattctgcg atgcggcgaa 4200
acattctata cacagthttc aaactacaaa aaatacaaac cthtagcctg thttatcaaat 4260
tagttagcta taaaatgccc atthttcttag caatatcgat caattgattg tcatctthca 4320
aagthtcaat aaaatthgtg gcagtatagt aatcccttct caatatgtcc aatthcttat 4380
cacagtcatc tataaattcg cgcttaatag attgcactth taggttcacc caattgttga 4440
atgtcatgat ccaactcttc tgcaggattt 4470

```

<210> 6

<211> 1151

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 6

Met Asp Gln Asp Asn Asn Asn His Asn Asp Ser Asn Arg Leu His Pro

1

5

10

15

Pro Asp Ile His Pro Asn Leu Gly Pro Gln Leu Trp Leu Asn Ser Ser
 20 25 30
 Gly Asp Phe Asp Asp Asn Asn Asn Asn Asn Asn Asn Asn Asn Asn Asn
 35 40 45
 Asn Ser Thr Arg Pro Gln Met Pro Ser Arg Thr Arg Glu Thr Ala Thr
 50 55 60
 Ser Glu Arg Asn Ala Ser Glu Val Arg Asp Ala Thr Leu Asn Asn Ile
 65 70 75 80
 Phe Arg Phe Asp Ser Ile Gln Arg Glu Thr Leu Leu Pro Thr Asn Asn
 85 90 95
 Gly Gln Pro Leu Asn Gln Asn Phe Ser Leu Thr Phe Gln Pro Gln Gln
 100 105 110
 Gln Thr Asn Ala Leu Asn Gly Ile Asp Ile Asn Thr Val Asn Thr Asn
 115 120 125
 Leu Met Asn Gly Val Asn Val Gln Ile Asp Gln Leu Asn Arg Leu Leu
 130 135 140
 Pro Asn Leu Pro Glu Glu Glu Arg Lys Gln Ile His Glu Phe Lys Leu
 145 150 155 160
 Ile Val Gly Lys Lys Ile Gln Glu Phe Leu Val Val Ile Glu Lys Arg
 165 170 175
 Arg Lys Lys Ile Leu Asn Glu Ile Glu Leu Asp Asn Leu Lys Leu Lys
 180 185 190
 Glu Leu Arg Ile Asp Asn Ser Pro Gln Ala Ile Ser Tyr Leu His Lys
 195 200 205
 Leu Gln Arg Met Arg Leu Arg Ala Leu Glu Thr Glu Asn Met Glu Ile
 210 215 220
 Arg Asn Leu Arg Leu Lys Ile Leu Thr Ile Ile Glu Glu Tyr Lys Lys
 225 230 235 240
 Ser Leu Tyr Ala Tyr Cys His Ser Lys Leu Arg Gly Gln Gln Val Glu
 245 250 255
 Asn Pro Thr Asp Asn Phe Ile Ile Trp Ile Asn Ser Ile Asp Thr Thr
 260 265 270
 Glu Ser Ser Asp Leu Lys Glu Gly Leu Gln Asp Leu Ser Arg Tyr Ser
 275 280 285
 Arg Gln Phe Ile Asn Asn Val Leu Ser Asn Pro Ser Asn Gln Asn Ile
 290 295 300
 Cys Thr Ser Val Thr Arg Arg Ser Pro Val Phe Ala Leu Asn Met Leu
 305 310 315 320
 Pro Ser Glu Ile Leu His Leu Ile Leu Asp Lys Leu Asn Gln Lys Tyr
 325 330 335

Asp Ile Val Lys Phe Leu Thr Val Ser Lys Leu Trp Ala Glu Ile Ile
 340 345 350
 Val Lys Ile Leu Tyr Tyr Arg Pro His Ile Asn Lys Lys Ser Gln Leu
 355 360 365
 Asp Leu Phe Leu Arg Thr Met Lys Leu Thr Ser Glu Glu Thr Val Phe
 370 375 380
 Asn Tyr Arg Leu Met Ile Lys Arg Leu Asn Phe Ser Phe Val Gly Asp
 385 390 395 400
 Tyr Met His Asp Thr Glu Leu Asn Tyr Phe Val Gly Cys Lys Asn Leu
 405 410 415
 Glu Arg Leu Thr Leu Val Phe Cys Lys His Ile Thr Ser Val Pro Ile
 420 425 430
 Ser Ala Val Leu Arg Gly Cys Lys Phe Leu Gln Ser Val Asp Ile Thr
 435 440 445
 Gly Ile Arg Asp Val Ser Asp Asp Val Phe Asp Thr Leu Ala Thr Tyr
 450 455 460
 Cys Pro Arg Val Gln Gly Phe Tyr Val Pro Gln Ala Arg Asn Val Thr
 465 470 475 480
 Phe Asp Ser Leu Arg Asn Phe Ile Val His Ser Pro Met Leu Lys Arg
 485 490 495
 Ile Lys Ile Thr Ala Asn Asn Asn Met Asn Asp Glu Leu Val Glu Leu
 500 505 510
 Leu Ala Asn Lys Cys Pro Leu Leu Val Glu Val Asp Ile Thr Leu Ser
 515 520 525
 Pro Asn Val Thr Asp Ser Ser Leu Leu Lys Leu Leu Thr Arg Leu Val
 530 535 540
 Gln Leu Arg Glu Phe Arg Ile Thr His Asn Thr Asn Ile Thr Asp Asn
 545 550 555 560
 Leu Phe Gln Glu Leu Ser Lys Val Val Asp Asp Met Pro Ser Leu Arg
 565 570 575
 Leu Ile Asp Leu Ser Gly Cys Glu Asn Ile Thr Asp Lys Thr Ile Glu
 580 585 590
 Ser Ile Val Asn Leu Ala Pro Lys Leu Arg Asn Val Phe Leu Gly Lys
 595 600 605
 Cys Ser Arg Ile Thr Asp Ala Ser Leu Phe Gln Leu Ser Lys Leu Gly
 610 615 620
 Lys Asn Leu Gln Thr Val His Phe Gly His Cys Phe Asn Ile Thr Asp
 625 630 635 640
 Asn Gly Val Arg Ala Leu Phe His Ser Cys Thr Arg Ile Gln Tyr Val
 645 650 655

Asp	Phe	Ala	Cys	Cys	Thr	Asn	Leu	Thr	Asn	Arg	Thr	Leu	Tyr	Glu	Leu	660	665	670	
Ala	Asp	Leu	Pro	Lys	Leu	Lys	Arg	Ile	Gly	Leu	Val	Lys	Cys	Thr	Gln	675	680	685	
Met	Thr	Asp	Glu	Gly	Leu	Leu	Asn	Met	Val	Ser	Leu	Arg	Gly	Arg	Asn	690	695	700	
Asp	Thr	Leu	Glu	Arg	Val	His	Leu	Ser	Tyr	Cys	Ser	Asn	Leu	Thr	Ile	705	710	715	720
Tyr	Pro	Ile	Tyr	Glu	Leu	Leu	Met	Ser	Cys	Pro	Arg	Leu	Ser	His	Leu	725	730	735	
Ser	Leu	Thr	Ala	Val	Pro	Ser	Phe	Leu	Arg	Pro	Asp	Ile	Thr	Met	Tyr	740	745	750	
Cys	Arg	Pro	Ala	Pro	Ser	Asp	Phe	Ser	Glu	Asn	Gln	Arg	Gln	Ile	Phe	755	760	765	
Cys	Val	Phe	Ser	Gly	Lys	Gly	Val	His	Lys	Leu	Arg	His	Tyr	Leu	Val	770	775	780	
Asn	Leu	Thr	Ser	Pro	Ala	Phe	Gly	Pro	His	Val	Asp	Val	Asn	Asp	Val	785	790	795	800
Leu	Thr	Lys	Tyr	Ile	Arg	Ser	Lys	Asn	Leu	Ile	Phe	Asn	Gly	Glu	Thr	805	810	815	
Leu	Glu	Asp	Ala	Leu	Arg	Arg	Ile	Ile	Thr	Asp	Leu	Asn	Gln	Asp	Ser	820	825	830	
Ala	Ala	Ile	Ile	Ala	Ala	Thr	Gly	Leu	Asn	Gln	Ile	Asn	Gly	Leu	Asn	835	840	845	
Asn	Asp	Phe	Leu	Phe	Gln	Asn	Ile	Asn	Phe	Glu	Arg	Ile	Asp	Glu	Val	850	855	860	
Phe	Ser	Trp	Tyr	Leu	Asn	Thr	Phe	Asp	Gly	Ile	Arg	Met	Ser	Ser	Glu	865	870	875	880
Glu	Val	Asn	Ser	Leu	Leu	Leu	Gln	Val	Asn	Lys	Thr	Phe	Cys	Glu	Asp	885	890	895	
Pro	Phe	Ser	Asp	Val	Asp	Asp	Asp	Gln	Asp	Tyr	Val	Val	Ala	Pro	Gly	900	905	910	
Val	Asn	Arg	Glu	Ile	Asn	Ser	Glu	Met	Cys	His	Ile	Val	Arg	Lys	Phe	915	920	925	
His	Glu	Leu	Asn	Asp	His	Ile	Asp	Asp	Phe	Glu	Val	Asn	Val	Ala	Ser	930	935	940	
Leu	Val	Arg	Val	Gln	Phe	Gln	Phe	Thr	Gly	Phe	Leu	Leu	His	Glu	Met	945	950	955	960

Thr Gln Thr Tyr Met Gln Met Ile Glu Leu Asn Arg Gln Ile Cys Leu
 965 970 975
 Val Gln Lys Thr Val Gln Glu Ser Gly Asn Ile Asp Tyr Gln Lys Gly
 980 985 990
 Leu Leu Ile Trp Arg Leu Leu Phe Ile Asp Lys Phe Ile Met Val Val
 995 1000 1005
 Gln Lys Tyr Lys Leu Ser Thr Val Val Leu Arg Leu Tyr Leu Lys Asp
 1010 1015 1020
 Asn Ile Thr Leu Leu Thr Arg Gln Arg Glu Leu Leu Ile Ala His Gln
 1025 1030 1035 1040
 Arg Ser Ala Trp Asn Asn Asn Asn Asp Asn Asp Ala Asn Arg Asn Ala
 1045 1050 1055
 Asn Asn Ile Val Asn Ile Val Ser Asp Ala Gly Ala Asn Asp Thr Ser
 1060 1065 1070
 Asn Asn Glu Thr Asn Asn Gly Asn Asp Asp Asn Glu Thr Glu Asn Pro
 1075 1080 1085
 Asn Phe Trp Arg Gln Phe Gly Asn Arg Met Gln Ile Ser Pro Asp Gln
 1090 1095 1100
 Met Arg Asn Leu Gln Met Gly Leu Arg Asn Gln Asn Met Val Arg Asn
 1105 1110 1115 1120
 Asn Asn Asn Asn Thr Ile Asp Glu Ser Met Pro Asp Thr Ala Ile Asp
 1125 1130 1135
 Ser Gln Met Asp Glu Ala Ser Gly Thr Pro Asp Glu Asp Met Leu
 1140 1145 1150

<210> 7
 <211> 33954
 <212> DNA
 <213> *Saccharomyces cerevisiae*

<400> 7
 tttcttggtg cttcaacgga ttatcttaaa aaaatctatc atatttcaaa atataaattc 60
 ttatttttac aaagaagata tagattatgc ataataattat tttgttacat tttttttctt 120
 ttacttttta ttttcttttc ttgtacttc ctcaaataag catcttttagc aggttaacaa 180
 tagcaaatat catgtccagc ccaaatacaca taatactttc cacttcactt tcactatctg 240
 tttcgggaga cctatcatcc acggtattga cagttggtct ttggaaatca gttgcatgtg 300
 cccaaggaat ggatctcgaa atgttctcaa acacggagga ctttctctg aaagtccac 360
 ggtaagcggg cataggttcg gcaaaactgt gaggattatt gtcatgtgaa ggcgtgatag 420
 gstatgtcgtt aagtgatgaa aaggcttttg aaccdaaatc actactatcc atagaatcat 480
 caccgacctg cgcaattgga taaggcagag gagcgtcacc ggtaaccctg aataggtgaa 540
 cattggcaac atgttcattt tgcactctga agatgatcca aacgaatctt cttaaaacct 600
 ccaaaagggc caatataaac gacgtcacag cactctgctg aattgtttga ggagcaattg 660
 cgtatacaat ccactcaaat ctaatgagga tatcccaaag catagcaaag taataaacta 720
 acttctact gaataataa ctgccgtttt cccaattttt tttaccagct aaatacaagt 780
 cgtctcttag caaccaattg taagaagtag tattgtgagc aaaggaccaa tccataacta 840
 aatcccaggc ggaagtaagg atcgaattca acgtagcgca cagataaaa ggggttctcc 900
 tttgttcaga acggtctgac aatctgtaag cacaagtgt ggcattatac gctataacca 960

aagtgtattt	cgccgcgttc	aaaagatgag	ggaaccaatc	accggaatca	gcaaattcttc	1020
gtaaaccattg	catgaatctc	caataacttg	gtaaaccatga	taaaacaccc	attgctctg	1080
aatgtgaaga	gccacataga	ttgttggcg	tatgagagta	aacacagaag	aacatggcaa	1140
tatcggaat	agaatacgtc	aacgaacaaa	taatatctcc	caaaaagaaa	tcaccaact	1200
caacaggga	aaaaccagac	atcatcagtc	taataagagt	aaccaccagc	cattttctag	1260
tgtgtacaac	tttgtccaa	tatggaatta	gaccagaggg	acataaaaaat	aagaaagaca	1320
caataccaat	atataaaaag	cctaagggtg	tgagcttctc	caaagcgaaa	cttaacatgg	1380
aacagaccgc	acaaggctact	atgaagaatg	tcaagaaata	caacttgagc	ggaattttac	1440
tcgtggcaaa	atcattgttg	aaaaattgcg	taccattctt	ggattgaatc	tcaccagca	1500
taataaatct	ataattaatc	ccggttctat	gccagataaa	acaattaaca	agaaagagaa	1560
acgcgatcaa	caaaaccatg	taccaaccac	cccacagtgg	gaacagaatc	ttgtgcgtaa	1620
atgacgtttc	ttcagagcta	atgcccaaat	acaaagtgtg	agtgatcaaa	gtcattgaca	1680
cacctatgcc	cagcccagaca	accagcatct	gcacaatcga	tctgttgttt	ctatgaacca	1740
tttgttcgga	gatagagtac	tgaatggtca	atttcttcaa	cttatgggta	ttgtgcttcc	1800
tgctcttagg	cgaattgggtc	aacgcagtgg	tgaaccactc	ggtgatctgc	gtttccaacc	1860
aagtgatggg	ctctttgtct	ctttgggcgg	aagagagctc	cgaagtgggt	gttgggtg	1920
aactgggtat	ttgttgcatc	ttctgtgcc	ccaactgcac	gttggcatcc	gcgtgcttga	1980
aaaggggtgta	gtgggtcctg	gcgtacgaca	tgaacgtggt	cagctctctt	gtgtgacacg	2040
ttttgtcgaa	tttcttgacc	attttacgga	acccgggtgac	gttgatatcc	cggaacgatt	2100
tgaccaattg	caagtacaga	tagtactcga	ttatagcgtt	gcttaacaag	tttctagcct	2160
gggttgctgt	catagtctcc	agaaaagaag	cgccaaaagc	aaacgtctcg	cggcctcttg	2220
aggacgcatt	ctgtctcaag	tcctgtagca	gagaaaatcc	acgtttgggc	cacgatggca	2280
gcagcctatt	gtccttaaga	actttctgac	agtaacgtag	cagcgacaaa	cgcgcttctt	2340
tggcctcttt	ggtgcagggc	attgacccat	acataacaga	ccggctatcg	ctgtcgatgc	2400
tattgacctt	ggagtacagaa	cgaccagcaa	gccctgccgc	atatagactc	gtggacatgt	2460
caacattcga	actagaccgg	ttcaagttgt	cgcgctcata	attcttctgc	agcgagtaat	2520
agtgcagttg	cgactgtaaa	acctcaaat	tcttgtcgca	ctcctttagc	aaccaaagg	2580
aaaactcggt	acacttactc	aactggaacg	agatcaacca	gtcttcaatg	aagtcagcaa	2640
cgaactccct	ctgcaaagca	gaatagctct	tcttgaacgc	aggccggag	cgatagtctc	2700
cgctcgctg	acttttgccg	ggctctctct	gctgaaatgc	agtctggtat	accgacacag	2760
agggcattcca	gctccggtag	gagctggatt	gctcttcttc	ggcatccagc	ttctccttgt	2820
agcggcgaag	cttcttcttg	ccgaccttat	aatcaatata	tttgtccctc	cattccggga	2880
tggcagactc	gggttagatgg	tcagcaaat	tcatttgata	cggtgctatg	gtaaagccct	2940
tcatttgctc	ttcaatgaat	taaaaaaaa	ctaccaacca	ctaccaacca	tcaaatcaaa	3000
caaaaacaaa	caaaaaagca	aaggagaaaa	cccatcttgt	acatgtacag	tcccaaactg	3060
ttgcgaaacc	gtgcgatgat	gttcatggaa	cttgcgctcaa	ataaaaccgc	atcccgctcc	3120
acgtgacaaa	cggcacccaa	aatccgttct	agaaatgcct	cacaccaca	ccgatgatgg	3180
gcttccgact	cgatgactac	caacacatag	ccagccgcag	cggctggcag	catgtgcatg	3240
ataatgataa	gtgatgatga	atgatggtga	taacgatttt	gaacaaatgg	ctgcacggat	3300
agcaactaag	gggatggaaa	gacagatctg	ggagagataa	ctgcaggggtg	tggcagggca	3360
acacaaggct	attgtattgc	actaaacggg	caagaagcca	tgatgtgcgt	ctgtatccca	3420
aaaaaaaaac	taatggattg	gcgcgtgtac	tatatatatt	catatgtcgt	gtgtttgtat	3480
atgtgtggga	gtgattgtgc	gtgtatatgt	gtgttggcgt	gtgtggtaca	atgtgtgtgt	3540
tttaatgtag	aaatgaggtt	gtagcacgtg	atcgggaagc	cacagtttgc	gcggagatat	3600
tttatttttt	ttcatcagcg	taagaagaaa	gcaaccttgc	agtctgtatc	gtaagagaag	3660
actgcagtta	aagaagttaa	gagaagaggc	ttgagtatcg	gtaaaggggt	gtgtgtttgg	3720
tgatttataa	aggagaaggg	catgaggaga	gagaggcaaa	ggatgatgag	tttcgaggac	3780
aaggacaagg	acgaccttga	caatagtaat	agtaataaca	gcagtgaat	gacagatacg	3840
gcgatgatgc	caccattaaa	gagattgctt	attacgggca	gtagcgatga	tttggcacia	3900
ggatcatcgg	gtaagaagaa	gatgacgatg	gcgacgaggt	cgccatcgct	atcaccgat	3960
ttggcgacaa	acgacagcgg	cactagggta	cagccattgc	cagaatataa	cttcaccaag	4020
ttctgctatc	ggcataaccc	ggacattcag	ttctaccaaa	ctcatacagc	gtgctacaag	4080
caggatttga	aacgaacgca	agagattaat	gctaatatcg	cgaagctacc	cctgcaggag	4140
caatccgaca	tcaccacat	tatctcgaag	tacagcaatt	ccaatgacaa	gatacggaag	4200
ctgattctgg	atgggatcct	atcgacgagt	tgcttccac	agctttccta	catttcgtca	4260
ctcgttacac	acatgatcaa	gatcgacttc	atcagcattc	tgccgcagga	gctgtcgctg	4320
aagatcttga	gttatctgga	ttgccaatct	tttgcacacg	ccacgagagt	gtgccgcaag	4380
tggcagaagc	tcgcgggatga	cgacagggtg	tggtagcaca	tgtgcgagca	gcacatagac	4440
aggaaatgtc	ccaactgtgg	ctgggggctg	cctcttttgc	acatgaaacg	tgccgcgata	4500
caacagaata	gtacaggatc	tagcagcaac	gcagatatcc	agacgcaaac	tacgcgacct	4560

tggaaagtca	tctacagaga	acggttcaaa	gtggagtcaa	actggagaaa	gggccactgc	4620
aggattcagg	aattcaaggg	ccacatggat	gggtgtgttaa	cgctccagtt	taactacagg	4680
cttttgttca	caggctcgta	cgactccacc	ataggatat	gggacttatt	cacggggaag	4740
ctaatacgaa	ggctcagcgg	ccattcggac	ggcgtcaaga	cattatattt	tgacgataga	4800
aagctgatta	cgggctcgct	cgacaagacg	atccgtgttt	ggaactacat	aaccggtgaa	4860
tgcattttcca	cgtatcgagg	ccactcggat	agcgtttctga	gcgtagattc	ataccagaag	4920
gttatcgttt	ccggcagtg	tgacaagacg	gtcaaggat	ggcacgtgga	gtccaggaca	4980
tgctacacct	tgagaggcca	cacggaatgg	gttaattg	tcaaattgca	tccgaaaagc	5040
ttttcatggt	ttagttgcag	tgacgatacc	acaatccgaa	tgtgggat	caggaccaat	5100
tcatgcctaa	aagtgttcag	gggtcatgta	gggcaggtgc	aaaagatcat	accgcttacc	5160
attaaggatg	tagaagatct	agccaccgac	aacacttctg	atggcagctc	tccgcaggat	5220
gacccaacaa	tgactgatgg	tgacagcgaa	tcagacacac	cgtcgaacga	gcaagaaact	5280
gtcttagatg	aaaacatacc	ttatccaaca	catctactat	cttgccgact	ggataacaca	5340
atcaaactat	gggacgtcaa	aaccggtaaa	tgcataagaa	cacagtttgg	gcacgtggaa	5400
gggtgtttggg	acatcgccgc	tgacaacttc	agaattataa	gtgggttctca	cgacgggaagc	5460
attaaggtct	gggacttgca	aagcgggaag	tgtatgcaca	cgttcaacgg	tcgaagacta	5520
caaagagaaa	ctcagcacac	acaaacacaa	tccttgggtg	ataaagtcgc	ccctatcgct	5580
tgtgtttgta	ttggagattc	agaatgcttt	agtgtgatg	aatttgggtg	cgtaaaaaatg	5640
tacaaattcg	atctcaatga	ttaggacctg	tgtgtggtct	tttcttggtc	aaaaacatcc	5700
gtagtacctc	gaatatatat	gcttacatat	ataatgaaaa	atacataata	gcatttttaca	5760
tttattttat	ttccaggaaa	aaaataaatc	gctagcgtat	cactagcctt	ttccatcttc	5820
agtttttttc	cttcctcttt	ttcgtctcga	taaaactctt	gaccattaga	tgctattact	5880
ccccacgga	gcccttccca	cgctaagggg	gatgtcaagg	gggggactga	agataaagct	5940
tatttatattg	acggcgatga	tgaagagggc	gtacgcataa	aatgtcgtac	aagatccata	6000
atctgcaata	ttgttccttc	aacagtacta	ataaatccaa	agggcgagg	tactttatta	6060
tataaacg	cacggccatt	gtatataaag	gcgattacat	aatagaatac	ttatcgattc	6120
cgtaatcctt	gaaatataga	atatacataa	aatagacaag	aaccaacag	ttgaggcgcc	6180
agctttggta	atagcttaat	aataaataaa	aaaagataat	atagctagct	agtattattg	6240
cttttatccg	tatggccacc	accacgcaac	cacaaaatat	actgatggat	gaaccttta	6300
atcttcctaa	taacagtgcc	cacaacaata	actatggaaa	cataaatg	aatataagaa	6360
cttttgctgg	tatgagcatg	cacatgcacc	ctgccaggct	gaactctctg	gagtttttgc	6420
acaagcccag	gagactatct	aatgtaaaac	tgcacagatt	acctcaggac	gagcttcaa	6480
gaaataacgga	catgaataag	ggaatgtatt	ttaatgggaa	acaagttcat	gcccattacc	6540
cgtttataaa	ttctggagcg	aactttaacg	cacatcatca	agacgtcagt	aaattaggcg	6600
aggaggaaga	cgaaatctct	cctctatcac	atgataattt	ccagtatgaa	tccgaggaaa	6660
atggtaatcc	ttcacctccc	atttacaaga	aatctggaga	actggttaag	agttcattaa	6720
aaagaagatc	caagtcctta	cccattactc	ccaaatctat	attcaacaaa	actggctcta	6780
agagtaaac	tgtcaattta	gatcatgtag	atactaggct	attgcaaaga	agtaaaagt	6840
tccatttcga	tcgtgtttta	ccaataaagc	tgttcaatga	aaatgagaaa	cccataagtg	6900
ttggcaacaa	aatggttcaa	caagatgttc	tgaatttcaa	gcataagcct	ttgacgagac	6960
taagtgcctt	taatggcggg	agtgatagcg	tacctataga	agatttacta	tctgaaaaca	7020
accaaaccga	atatggagat	acatggctac	aaaaccggaa	gggtgtattc	ctattttggt	7080
caaatttctaa	taatcgtaga	aataaaaaaa	agaagtttaa	actaagtgac	gatgacagtg	7140
atattgaaaa	tgacaatgat	agtgacgacg	ctataaaccg	tttagtaagg	caacaagaca	7200
aagaccaagc	tcattcttgcg	cacgggttga	agaatttgtt	aataaacgat	gacgacgatt	7260
atttagaaac	aagaacaaat	tctgtctaat	caggagccaa	cttgtttatt	ggaaactcta	7320
aaagaattgt	tggtctttac	aacaagaact	ttccaatatt	aagtgcagg	aaccgtaaga	7380
gtttaaaact	taacatattt	ctgaatcttt	cccgtggaag	gcccgttttt	ttgcaagaaa	7440
ttacgctact	aactggcttc	cacaacatgg	ttataattgg	caaagtcttt	gtgaaaaaca	7500
tatactttga	taagaagatt	atcgtaagat	atacatggga	tgcatggaga	acttttcatg	7560
aatcagaatg	cgtgtatttt	tctaattgcca	atggcatctt	accaggaagc	aatatggata	7620
tttttaagtt	ctccattgat	gatatacaca	atccaaatga	taaagatagc	aatatatcac	7680
aattggagtt	ttgtattcaa	tacttaactt	ggggcggtga	togttctagg	aaggaatatt	7740
gggacaataa	tgattcggca	aactataaaa	ttgacgtggt	aacgaatgaa	acgaggacag	7800
ggcccacaac	agacgtcaat	gataactacg	agatgaacaa	tagtcttttc	agaaacccat	7860
tccattaaag	aaaaaaacaa	aaactaatag	tattgtttta	catatacaaa	agaaaaaaa	7920
aagataagaa	tacgtacttt	ctattttttat	ttgaaggtca	ttttctttat	ataaaaataat	7980
atactggaat	gatgttcttt	ttgttgcttt	ctatatttgc	ttttcttttt	tactttttat	8040
tcaactccag	acgttcttaa	ataactcggt	atcatccaat	gaaatgttag	aactggaatt	8100

tgaattcata	ctattattat	ttgtagcgg	aattccagtg	ttcacgaatg	gtgacggagt	8160
attacaaggc	tgaggtgtcg	tggaagtga	aaatgagttc	tggctttgag	tctgggcaga	8220
aggttttgag	tataaagata	aaatggat	tttcaaatac	tccctttgcc	caaattcacc	8280
gactctgttg	ccagtccttg	tattagcagc	acctatgctt	gttgcagaac	tagttacact	8340
tgtattagat	gtagtcttag	aaagcagatg	aaactgtagg	tttaataaag	tcgtattact	8400
ccttgaggtg	cttaataagt	gggaaggtgg	ttgtgtttgc	gtctttttgt	atgaatttga	8460
gctctggtcc	aatcttgcac	tcgaagcagg	caacgaatga	ttcgcagacg	gcttatgcag	8520
aactgggtca	gtagaatcat	tttaacccttc	aatactggag	aggtcaccaa	tccacttctt	8580
atattcatat	ttattcttga	tgaaattttg	taaactgctc	gtatcagtga	ttttctctcg	8640
tttcaattca	tcagccaaag	tagcttcgta	gtatgagtta	gctcttaaat	tgtttttgaa	8700
ctgtatcagt	ttcaccaaat	gttcttcctt	ccatgtatcc	agatcaacag	atttcacttt	8760
cgaaatgtgc	gttcccaatg	atctatgtat	accagcacat	ttaatgcaaa	tgaaaaacacc	8820
aagtgaccag	gaagcccagc	gtggatgtag	ttgcgcctta	cagtcggcac	aatgactgtt	8880
tcctggatcg	cgtaaaagag	cacttaatgc	cttcttgact	gggactgacg	tcgacatcgc	8940
ttgtaaattg	ccccacgttc	tatgctcctt	tgtacagttg	aagcaagtat	tacctttact	9000
tttgttacta	ttagtaccga	tcttatacgt	attatcgagt	cgaaaatcgt	ccagcacgaa	9060
atatggaaac	ggtgcgttta	ggtgactaaa	ctaaatagcg	tatatattat	atatgttcca	9120
atctatatat	atatatatgg	cactctgtta	cagttcctta	tcttaaaaca	caaacacctg	9180
gtcttccatc	ttggaaagcg	gtttggaacg	tctgaacccc	aagtccacgg	tacttctctt	9240
aactgaggga	accatggctg	gaggaccaca	gatcaaaatt	tgaacgttgt	ccattgtagc	9300
agcgggcaag	tgttccttga	tgacatcctt	ggtaatgtat	cctacaccac	cagtcagctc	9360
ttcacggtca	ggagagtcta	agtagtaaac	tatcttaaat	tgggaaggct	tcattggcac	9420
caacgcttcc	agttccttct	tcaacagaat	atcctcctca	tggacgttcc	caaagactag	9480
agagaccttg	gtagtgtcgt	gaggggccat	ggcaatagct	ttcatgatct	gatacatggg	9540
cgcaatacca	gtaccaccag	caatcatccc	tagatgggaa	cggcagtttc	tctcataatg	9600
atagttccca	cgagggccct	tgatctggat	cgagtcaact	atcttcaact	ctccaatcat	9660
cttagaaaac	ttacctgtgg	ggtaagactt	cactagtaat	tcaaagtttc	cctttgtatc	9720
ttcatccaac	gatgtgggcg	tatacgatct	ggtaatatcc	ttaccattga	tattggcctt	9780
aattacgata	tgctgaccaa	ttggtaaaac	gagtaacgtc	tcagcatgag	gtagcccga	9840
cttgtagatc	gaagtattat	gcgttaagat	ggttttttca	accagcggaa	atgattggaa	9900
gtcattcctt	ttgggatcca	gcacaggctt	ggctctcggt	ccgataatga	acttgaagag	9960
caaaggcacg	accacgatca	cgatgaccac	cacaagcttt	tgagcatcaa	tagccatctt	10020
cgattgtcta	attgattgag	tctcttggtt	aagagccagt	tgaatgcata	ctttcaacac	10080
ttttttttct	cgttcatttt	ttcgctttat	ataataagaa	tatttatata	tatgtagata	10140
atttaaagaa	gtagaaagag	cccttaatgg	taccgcttta	cgaggaaatt	tagcactgtt	10200
acttttttga	gagtagtgca	gtcttagaag	ggcctttgag	cttgatgtac	acgtccgtgc	10260
cccagccaag	taaactttga	acgtcaatct	tgcccccaaa	cagttctagg	tacgtcttgc	10320
acattggtag	accaaagccc	atgcccggaga	cgttggtgat	ttgctcgccg	ggcagatctg	10380
tagactcaga	gtcagctgat	tggtgggtat	gtgtagagta	agagtagttg	aacatggagg	10440
cctctacctc	gggcgtgatc	ccaccaccat	ggctccgtat	ccgtaagtac	agctcgtcat	10500
cgtctggctt	taaaagattg	atctctatgg	gcatatgttc	ctttccaaga	gctatctgag	10560
cttcaaaagc	gttcttgaat	acctctgtca	taataacttc	cagaatgggc	ggaatgcacg	10620
tgaatgtgat	gtcttgtgat	ggcgggtgaa	ttaggaccgg	ggctccgtga	gtgttgaact	10680
tcacaaagca	tatatcattg	acataatcag	aaacgtgctt	gatcaattgt	gctatgggga	10740
ggtcccgtg	caagattcca	atcatccttt	tgttcgtgtc	acctttgttc	tgtgccatca	10800
acgacaaata	gtgcgtcaca	agcagcttca	tcgtgattct	ctccttcagg	tggaagttca	10860
agaactgcga	aatttgaaac	tttgggtaac	acgattggat	ttcctgaaga	ccttttgcca	10920
ataccacaat	ggcgtcctcg	tggtcgtcca	gcagtctctg	gaacttggcc	tgtatcttgg	10980
gaggattgtg	caactcgtac	gggtatgcga	ttgatagcaa	cgtttgcaag	cttttcaaat	11040
acaaagagtt	ggtctctcta	atatgagggg	tgattaccgc	attatacggg	agcctctgaa	11100
tggtcattaag	ccgttttacac	gtcaatgaaa	gcaacagggt	gattgtcttg	atcgtcagca	11160
tgtactcctc	cttcttggtc	agcggcggcc	tgtattgcag	gaaataactcg	tagtttaggg	11220
gcgcaattgg	tttactagcg	tagtctgtga	tcagcagctc	aatgttcgac	ctaactctgt	11280
aatgctggtc	gaacgacagt	tgcgacagca	gctcgtgtga	gggtcgctgg	cgatgtgccc	11340
agcgcattcc	cccacatttc	catgaacgca	taatcttcca	catactaccg	tgaaggcgct	11400
ttgataccac	gtcttccagt	tatggtcttg	actcaccaat	gttctcagtt	taatgttgtt	11460
tgcttcatac	tggcagacat	tttcccttat	taccgtgaga	agcgagcggt	ggattaatcg	11520
ggatgtcaaa	aacaagaaaa	ttgcaacctt	cttcaacctat	aacctatatg	attttagtagc	11580
agaggagtgg	aataagcaac	aaatcatagt	catcaatgtc	gtttaatgcc	ttcgccagct	11640
ctttgagcaa	gaaattacag	gaaatatcca	caagtgtttc	cgaaaaaac	cagggaattgc	11700

ccagtttggc	gcaatccacc	caaagaatgg	tccaggaacg	tttgggcccag	gtgacggaca	11760
tctcccaatt	gccaaagagag	tacacagagc	tgggaagataa	gggttgacacc	atcaaaactga	11820
tttacaacca	cttcttgggt	gtaactgcta	tctacgaaaa	cggatcgtac	gattacccta	11880
aatacatcaa	cgaatccgct	aacgagtttt	caagaagcgt	ggcttccaag	ttgacagagt	11940
tgactcatgc	tacatctgcg	tctgaggcac	aaaacatctt	agttgctcca	ggcccatca	12000
aagaacccaa	gacgttgaac	tacgccctca	gtaaagtggc	tttgaactcc	agtgaatgtt	12060
tgaacaagat	gttccccacc	gaggaacaac	ccttggcttc	ggcactcttg	caattcagtg	12120
atgtgcaggc	taagattgct	caagctagaa	ttcaacaaga	taccttgatt	caaaccaaat	12180
tcaataaaaa	tttgagggaa	aggtctctct	ttgagatcgg	taaggccgat	aagtgcgcga	12240
aggatgtcca	ttccatgaga	ttaagatatg	acgtggcaag	aactaacttg	gcaaacaaca	12300
agaagccaga	aaaggaagct	tccttgagag	tccaaatgga	gactttggaa	gaccagtgtg	12360
ctcaagtcac	tgaagacgct	actgtgtgct	tgcaggagggt	tatttctcac	gctaacttca	12420
gtgaagattt	gaaggaattg	gccaaaggctc	aagctgagta	ttttgaaacc	tcggctggcc	12480
taatgaaaga	gttcctatcc	aactcatttg	cggaagagcc	ggaagcaaag	cctgaggttg	12540
cagaagaaga	aaagccacag	acagctatct	ctatgaatga	cgaagacgac	gcttaatagc	12600
cgccctgttc	ttgtctttct	ctctccttct	ctatatatat	atatttactt	aatacaatat	12660
catattctct	agtgtcgtac	gaatattatc	atttatttgc	tggcttttgc	agtgaagtata	12720
ttttttgtag	agaaaaatata	attactggaa	gagaaaaagc	tactgaaaaa	aatttcgtaa	12780
cagtcaccaa	agggcaaatg	gagcaaccca	aaaatggacg	ctactcaacc	gcaatacga	12840
ctatccgtag	taacacaatg	cctaaagtct	gccattgacg	tcattccaatg	gttgattcct	12900
actattacta	agttcagcca	atcacaccgg	ttagttttcc	aactattgtt	tattttcttc	12960
acattttatg	tcttttataa	gttgttgatg	aacttcatta	ctctgggtcaa	aagattcctg	13020
tatttgcagt	tggtagtaac	atgtatcggt	atttatatgc	gtgggtcgca	gcagtttttg	13080
actgtggacc	tgttgaattt	ctacaatttc	gtcatgtcaa	atagatacta	cgcattttaa	13140
atttatactc	tgtttattaa	tgccctggaa	agagaaatca	acactgttta	tcatttggcg	13200
cagatgaaaa	tggaaacagtt	gcttaaatag	agcgaaacgc	ctctgcttac	tcggaacatc	13260
ttaaagtaat	tacatatctt	aatttctaaa	tattcatata	gataaagaac	gttttttttg	13320
ctagattcct	cccctagtca	ctatcaaatg	tttgagctatt	tttggtgact	tggtttttcc	13380
aaattattca	ttacaattgt	gacttgattt	ctgaagtttt	aagaaaaatt	gcaaaactaag	13440
ttattacaag	cgatttatcc	ggtcgcgatt	gaagaactga	aaacaacagc	agcagcattg	13500
taccaagaat	ccaagatgc	tcaggtgtgc	ggtcaagaag	tttgcattat	ttgccacctt	13560
cctcactatc	gtcgctaata	tttacatcta	tacttatcca	tctttccatc	cagaacagtg	13620
ttcatggaac	tgctctaaca	agaacgcacc	actacagaaa	gatttgactt	ttgtggacaa	13680
agtgaagaat	tacttcagtg	atgttcgaga	gcaatggcat	ggtagccatg	cctctgcagg	13740
taacgatgag	gatattcata	ttttggcctt	cggtagcct	cagattaaag	gcatctggcc	13800
aaagacgcca	tatgtgtccc	gattagatac	atacggtaat	gactactatc	ttggtcacat	13860
ctatgatatg	atgcagcaga	gactgaagcc	tcaggtgggtg	actgttatgg	gtgatctttt	13920
ctccagtcga	tggatcgggtg	attccgaatt	ccataataga	acgaaaagggt	atattagcag	13980
aattttcaaa	agtagcccaa	cttctattga	aaatatcaaa	caacagaacc	tcgatgagaa	14040
aggccaatat	aaggcgaatt	ggcctgaatg	gggagaccgt	tttaatgaga	ttttagacaa	14100
tgtcaaggaa	aacgaggctg	ataatcaaga	actttctttt	gggtttgggt	atgaaaatat	14160
ccactcatgg	aatccagact	tgggaagactt	tttaatcatc	aacatcacag	gtaatcatga	14220
tgtaggttac	tcaggcgacg	cgacttatca	acatatgaca	aggttccacg	accttttcgg	14280
taaggacaac	tattggattg	aatacgaaac	caacactacc	catccttgga	gaattgtagt	14340
gttgaatgat	cttttactag	aaggctccgc	tttacaacca	gaatttggtg	aagctacttg	14400
gatcttttta	aaccaactaa	atgaacgtaa	attcaatgggt	agtactgttt	tgttgacgca	14460
tgttccattt	tacaagcgtg	aagggttatg	cgttgatggc	ccagacacta	ggtattatcc	14520
agatgcccat	gccccagaat	catataaatc	tgggcttttg	agatcccaaa	atcatttgag	14580
tgaatccgtt	tcaaatcaag	ttttgaatat	gatatttgaa	aatgggaagc	caggtatcat	14640
attaaatggg	catgaccatg	agggtgtgta	aacagtttat	aataaaaaat	ctacatctac	14700
ttgggaagcg	actaaaaata	tcgaaagtga	cgttttcggt	aaagaaatca	cagtttaagtc	14760
tatgatgggt	gaatttaatg	ggaatactgg	tttggttaca	ggacacttca	acacagattc	14820
aatgacttgg	gagtggactt	tcagtttatg	cccattcgct	attcaacatg	tttgggtgggt	14880
cgctaaagtg	tccttgcctg	tcaccatatt	cacttgggtcc	tcgctactat	ttgtctaaac	14940
aaaaaagagt	aaagaagata	aaactcctgt	atagagattt	atattaaatg	tatgtatata	15000
tatgcatata	ctgctcttca	aattgcttat	gttttataaa	cattgacatc	ggggttccac	15060
acatcaagat	aatcattata	ataccttaga	ttgtataaatg	tttacttact	actttaaaaat	15120
aactataaat	aaaattaaaa	atcaaattaa	aaatcaaat	atactgcat	tagcactgat	15180
ttttgttgtt	tttgttttaa	agggatgcac	attttttttt	tttatgttat	tatcatcac	15240

taattttttcc	ctgttttcaat	tgttttcaaga	gctgcttgcc	gtgagaaatt	tcttgttttt	15300
gttgtttctaa	agttacatta	ctctgatcat	taacattggt	attgtcaaca	ctcacatcac	15360
ggacctcaga	cgctgtatct	accggtggct	gcaaaaatga	ataatccagt	ctgaagttaa	15420
ttttatcaat	tactgtccaa	tcgtccagtt	taaaaatttt	gtaatcacca	acttcgcaa	15480
tttcattaaa	aaacttgacc	tcaccttgtc	ttaaaaacca	cattgtacca	tcctttgaaa	15540
ctttccaatc	cctctcattt	agaatcttgt	atgcaatctc	cctttctaaa	ggagtaatag	15600
caaaataata	gttataaaat	aacgaaaaca	tttctaaagt	tctgaaattt	tctagaattt	15660
tggcataaat	tgacgatgaa	tcttctttca	gcctttcgga	gcctataata	atatcacgca	15720
acgaacaacg	cataacgtcc	cactgttgag	tagatctgaa	tgcatccaag	ggagatgggtg	15780
gattaactcc	ttgcggtact	tcgaccagcc	tggaaatttc	acacatatca	cgagaccttt	15840
tgtacgttat	tttcgattca	atttgactat	atgagagttc	agaactcatt	atgaattctt	15900
gcacaccact	tggtagaagt	aaagtttcaa	aatcacttac	aaattctttt	ttgatttcat	15960
ttgtcttagc	ttctctttgt	tccaagctta	aatatttcgg	tgtgcttggt	tgttcttctt	16020
cctctggctc	cgtctctgaa	tccttatccg	attcaaaatc	gtcatcgaat	actccaaaag	16080
ttggcacttg	caatttattt	ctttcttctg	gtgactcttc	ctccttctgc	tccataacct	16140
tcttagttgt	gttctcaggc	ccattttggt	ttaccgtagt	tgcaacttca	tgttcagact	16200
ttggatttgt	cgccaaggaa	actaaaggat	gatgtgagga	gctcatattt	ctatagaagg	16260
cctgattgtt	ttgatgcaca	gcagccgccc	cagcagccaa	aacagcagcg	gccgttgctg	16320
ctgaagatga	tgccgaggta	ttgtcaggtt	gtaaagacaa	agatgaagta	gataacttag	16380
gtgtattcaa	cgcactacca	atcctgctgt	tagcgttgga	tgatgttgta	gtagcagcag	16440
cagttgttgg	agtcttttgt	gaggtattgg	atatgggtga	cgatgcagaa	gtgacttttc	16500
tatccttttc	taccgcttgt	gatgcagcaa	cggcccatct	taattcacca	gcaggttttg	16560
ctggaagagt	agcaggtttc	aatgtagtgg	caccagtaat	accgtttgga	gtttgatgaa	16620
tatgtgtatg	tacattttgt	ggagtttctg	gtgtagctga	acttggtgat	ttttgcaatg	16680
acggaagcaa	attatcagca	ctgctcctgg	gactctttat	agaagtttta	acagcctctt	16740
caggctttgt	tgcattatgg	ataggggaag	atgaaggact	tctttccggt	tcttttgaa	16800
catcagcaac	gggaatgaca	ggtgacggcg	tagaagaggg	accagcaacg	ggtatagctg	16860
cacctgttgc	attttttgcc	gcaagtttag	cagccttttt	ggcttctctt	tccaattttc	16920
tttgctcctt	tttagaaagt	tttgatatgt	cctgcagaga	ttcattagca	tccgatgtat	16980
tgttatcttc	agcattttgt	gaagcaaaat	actgggccac	ctcatgagca	atggcttctg	17040
tactctgtaa	attcaaaccg	tcataaatag	tttcatcctc	tacgaaatct	ggatcctgat	17100
ttgactccac	gaagtaattt	atatcgctct	gcacgttttt	aacatcttgg	gggtccaatt	17160
cctcattggc	taacaatctc	aaggctaact	ccatctgctg	ttgatgccac	ctatacctcg	17220
cttgaaaacg	cttatattgt	tccttttttt	catcatttgt	cgtagatgac	gttttctctt	17280
ttttattaag	aagtaaaagc	ttgtcaattt	ccacttgtag	ggaatcgat	tgcttttcca	17340
gtcatcgat	catttgggaa	aggtactcag	atatatctct	cctttccctt	tcttgccggg	17400
ctaaagtctc	tgattttttc	aggctaatat	tagagtatgc	tttttctttg	gacgccttct	17460
ctacagcctt	atactttttc	atcgctatatt	ctacagacct	tctgtaatct	agaagagaa	17520
ctttatcttt	aatatcggtt	gagcttttgc	atgattttat	ctgttccctt	agcctttgca	17580
gctttttgac	ttctcttttt	aaatccgact	ctagcttggtc	cttttgggaa	ggattgtttg	17640
tgcatgattc	atgtctttcg	taatagctat	tgaagatttc	taaaccttcg	ttaatttttt	17700
taaagaccct	atcgacctcc	tgctgtaatt	ttctatgagc	catgtcgatc	gactaaacat	17760
acaatgaatc	ttcaaaaatc	taaatttgaa	aatgaataaa	taaatgtgct	gtagataata	17820
gctttttgatt	ttttgtcttg	acacaatgag	acttctcggt	aaagtcaaac	caaaaggaga	17880
tgaacaaaag	aaagttcgaa	atatcaacaa	aatgctttta	aaaatgtaga	agcaaagtta	17940
ttaataactt	tacaagcgcc	ttagttagtt	tattgatctg	taggtcagta	ggttttagag	18000
attacgcaca	gatgaaagtg	aattaagcta	ctattacata	aaatcaaggt	atctttcaca	18060
agtttttcac	tcagcgaaaa	ttgcgagacc	gctcgtggaa	ggtgacgata	atacgatact	18120
tacaacttta	atattataaa	tataagcatg	tgttggaaca	atatagtcag	gcaaactgct	18180
aaaagtcccc	acttataaac	agtgtgtaaa	aagtaacgtt	gtttttgctc	cttatccttt	18240
atttggcggt	ttgacagtga	agacaatgcc	gctaccgatg	cgttggttat	attaattcct	18300
cgattaattta	taccatatatt	ttgaatgtcc	agtgcacaatg	aagaatttgt	attcgtaact	18360
tcgtatgtcg	gtaggaaaata	ctgaccaggg	aactcattaa	ttgaatcgct	gcctaacaga	18420
ggctcaacct	gcaggtctat	cccttgccaa	ttttgagaac	tctcattaag	atatttttca	18480
aactgagtg	ttgtgtcggt	aaccctaaaga	tttaggatgt	ccgttgattg	cactttcaaa	18540
ttggggtcag	tttgaaggct	taaatttgct	gtaatgaact	tttcattcac	ttccatttgg	18600
aaattttcac	acaataatga	tacgctgtca	ttaatgaaac	cctctatcgt	agttgttata	18660
tttccatctg	tccaaaacttc	ccttttgat	agcggaaggtc	ccgctctttt	gtggaatgta	18720
ggcaaaactgg	gtacttcaaa	gtttgttaat	cgcgtcacta	tctgccaatg	aattaataac	18780
atcaaaaaaac	caaacagcca	caggtgagcc	ccatttgaga	ttatccacca	gttacagtaa	18840

aatagttttg	atTTTTtgac	ttttttgtca	tcgatttttt	ctggttctct	tttaaatt	18900
ttttgattga	taagctttgt	taattggtag	ggaattggat	aaagtgtact	gagattcaga	18960
gtagtaatta	agcttttcaa	aactttgttg	tatttagtaa	aattggtttg	tctgtcagg	19020
ccattaattt	gggacataac	caggtttact	tgctgatttt	ccagctgaaa	cagaattctt	19080
tcgattgcc	tgagcaaaat	cacaaatgca	aagtacatga	ctgtgaaaat	tattgtcatt	19140
ctcttacatt	tacggaaaac	gtgatttttt	actgcactta	gtacactggt	ttcctcgcta	19200
tttgttttca	gagaccgttt	aattattgaa	gtagacgatg	taacctcggt	actccccca	19260
gataaaatgg	aaatttcatt	ctgtaagttt	gcagtaaaagt	tccttttcaa	ctgaagccag	19320
tcagtggaa	ttcgtagaaa	attcaaaact	aaggtagtag	aggtatctaa	ttgagtagaa	19380
aagtcagtga	tagtgccatt	gaagtctctg	aatgatttca	atttccttga	gtagttccca	19440
aaaatactgt	caaacatgga	gtaatttaac	agtaggtttg	atgccatttg	tgaaaatggaa	19500
tccttatcaa	tggtcattgc	gtcttctgat	atttcattcc	atgtgtcatt	tatcagctgt	19560
gatttagtga	gtagttcggt	aaaaattgta	ttgtttacct	caagaacttc	cttaatcaat	19620
tcagtttcca	aacttatttg	gtcgttgata	agtttttgtt	tattatccag	tatgtcgtct	19680
gtgtatgact	tgaatttttc	ttgcatatcc	gtacctaccg	tagattcaat	actccgtagt	19740
ttgctttcga	aatactgatc	aattgtatta	tcaatgaatg	ttttatttaa	agagtaactg	19800
ttggcagaat	atagggcatg	ctgtgttggt	actgctgtca	ttgtggccgt	tggtgtggcc	19860
gtgtgtgtgg	ctgtcgctgt	agcagttgaa	atagatgcaa	aggttgagac	ggatcatttt	19920
cgtttgacct	gtgcgtgaga	catttttgcc	attgtaatgt	caatgactga	aaaggaaccc	19980
caaatagtta	aaataacaga	agtaaaagaat	atcagtaata	gaggatggaa	aatacaacta	20040
aaaaacctct	gtggaagtga	taaaggttta	attatgtgta	cattgttcat	ctttgccgta	20100
ctcttagttt	tatcgaaaag	tggtgtagat	tttcagctac	ttaagccatg	ttagtgttct	20160
agttttttgc	tttcatgtgt	atataaagg	ccaagatagc	atatgtaaat	gttttattac	20220
tggtgtgaaa	cattttttaa	acatttttca	ataatgtgct	gttataaccg	tacaaggacg	20280
gttataattc	tgaattaaaa	tgctgcaccg	aatgaaaata	ccaatgttct	tccaaaccta	20340
ttggcgaaaa	ttaagataag	ctgtcgccctc	tataatcaag	aataatctca	actattataa	20400
atattattcc	cagaagttga	aattgtgaca	catttgagta	atatcacctt	acacggaaca	20460
tagttgcccc	cttctccgaa	attgctgaca	atgataatcg	gggattcaat	gacaatgaca	20520
atgaaggcaa	aagattttca	tattgtactc	ttgttttttg	agttatcact	ttcaagtacc	20580
ttcttggtta	cagaaataga	accgctatag	aggttatcgt	gtatacgtat	ctgttggtgt	20640
ttctaaatca	aaaactcgct	tttaaagcag	tctttcaatt	ggaagattca	aaaggatcta	20700
aaaagatcta	aagtagtaaa	gaaaaaagta	taagcccaca	ccttttttgt	aggataatgt	20760
ttactgggtc	ggagtatcat	tccgtagact	ctaattccaa	caagcaaaaa	gacaacaata	20820
aacgtgggtat	tgatgacaca	tcaaagatct	tgaataataa	gataccgcac	tctgttagtg	20880
atacttctgc	cgccgccacc	accacttcta	ctatgaacaa	ttctgcttta	agtagattcct	20940
tagatcctac	tgacataaac	tatagcacia	atatggctgg	tgtgggtgac	caaatacatg	21000
attatactac	ttccaataga	aattcttttaa	ccccacaata	ttctattgca	gctggaaaacg	21060
tcaattcgca	tgatcggtgt	gttaaacccta	gcgccaattc	aaactatcag	caggctgcat	21120
accttcgaca	acagcaacag	caggatcagc	gacaacagtc	accctctatg	aaaactgaag	21180
aggaatccca	actctacggt	gatattctga	tgaattctgg	tgtcgtagag	gatatgcatc	21240
agaatctggc	cactcataca	aatctgagcc	aactgtcgtc	taccgtaag	tccgctccga	21300
atgattctac	tacagccccg	actaatgcgt	ccaacatcgc	caatacggct	tctgtgaaca	21360
agcagatgta	tttcatgaac	atgaatatga	ataacaaccc	acatgccttg	aacgatccat	21420
ccatcctgga	aacattgtcg	ccattttttc	aaccttttgg	tggtgatgta	gcacatttac	21480
ctatgacgaa	tccaccaatt	ttccaaagtt	ctttgcctgg	atgcgatgag	ccaattagaa	21540
gaagaagaat	atcaatctct	aacgggtcaaa	taagccagct	aggcgaagat	attgaaactt	21600
tggaaaacct	gcacaacaca	cagccgcccc	cgatgcccaa	ttttcacaat	tataatggtc	21660
tgagccaaac	taggaatgta	tcaaacaagc	cggctctcaa	ccaagcagtg	ccggttagta	21720
gtattccaca	atacaatgca	aaaaaagtta	ttaatccac	gaaggactcc	gcattgggtg	21780
atcagagcgt	tatttactcg	aaaagtcagc	agcgaaattt	tgtaaaccgc	ccatcaaaga	21840
atactccagc	ggagagtata	agtgatttgg	aaggcatgac	gacgtttgcy	ccaactactg	21900
gaggtgaaaa	taggggcaaa	tctgcactta	gggaatctca	ctctaactct	agcttcactc	21960
caaaatctca	aggatctcat	ttaaatttag	cggcgaaac	acagggaaat	ccaatccctg	22020
gtactacggc	atggaagaga	gcaagattgt	tagaaagaaa	tgaattgca	gcttcgaaat	22080
gtagacaaa	gaaaaaggtt	gcgcagctgc	agctccaaaa	ggaatttaac	gaaattaaag	22140
acgagaatag	aaatttactg	aaaaagttta	attactatga	aaaactaatc	tctaaattca	22200
agaaattctc	caaaatttcat	ttacgtgaac	atgaaaaaact	aaataaagac	tcagataata	22260
atgttaatgg	cactaatagt	agcaacaaaa	atgaaagcat	gactgtggat	tcattaaaga	22320
tcattgaaga	acttttaagt	atcgattcag	acgttacaga	agtgataaaa	gatactggta	22380

agatcatagc	catcaagcac	gagccatact	ctcaacgttt	cggaagcgat	actgacgatg	22440
acgatataga	tctcaagccc	gtagaagggtg	gtaaggatcc	agacaaccaa	tcattacceca	22500
attctgaaaa	gataaaataa	caaagttttg	gtgagcacaa	cggtgtttca	ttcaagaatg	22560
ataaggacaa	ctatatatgt	aatatacgtg	tataaatact	cgagctgcat	cctttctttt	22620
ctgccgttat	tatttcgttt	agttactaca	tttatgggat	aaattacctt	taagacatta	22680
ttatttttgca	tgtatgaaat	ttgctcgtgt	caacccctta	gtcctccttg	cctctgttgt	22740
atggaggaag	ccctttaaac	ctgtatttag	ttttttaaat	ctattagaat	taagtacaat	22800
tgtacagatg	gtaaaaaaa	gtaatcgtta	tatcgtttgt	cagtgatttt	ttttttttta	22860
tttcattcat	tatttatctc	ttatttttca	atttgttccc	ttattggggc	aaaccacggg	22920
tgtcccatcg	cttccttagc	agttaatctt	tcttgatgat	catatctcaa	aagattgtca	22980
ataaggtcaa	taatttcato	gttgccgctt	aaatgtttat	taccatcatt	gatgaatcta	23040
tgccaaggct	ttctgatgta	ttgggtccata	tcgtaaaatt	ctctcggtaa	ggtaatttca	23100
tacttcaaca	ggtatttctc	aaaatcgctt	gtaccaagta	ctttgacgat	cttgacaagc	23160
tggtctgtgt	tactcgttcc	atggaaaaat	ggctctcttt	taaagatcat	agaagccaac	23220
attgtcccaa	acgaccacaa	gtctaaagaa	taatcataca	ttctgtagtc	aactagtagt	23280
tcaggagcct	taaaaaacct	cgacgcaaca	cgaacattgt	attccatatt	aacatgataa	23340
aattcagcaa	gcccccaatc	tatcaatcgc	aattttttat	tcttatgatc	aatcattacg	23400
ttatgaggtt	taacatctct	atgcattatt	cccattgaat	gacaatagtc	taaggccttt	23460
aataactcaa	acataataaa	cctaatttctg	agatcgggta	atttagggta	aagaatacgg	23520
aagtcacat	tatctacgta	ttcgaagacc	aaagccggag	tttttgagat	gggatccttt	23580
attatatcaa	atagatgaat	tatgtttgca	tgaccattgt	gtggctgatc	aaagatatag	23640
ggcctgatga	attttaaaac	gtcctccttt	tattttgtgt	aatattgatc	ttttgaaat	23700
ggcaaagtgc	ttggaggcac	tttttcgtta	gacaaatccg	ttaagatttt	gatttctctc	23760
ttgatcttct	tctttttaac	tggtttcaac	atcttaataa	caattttaac	tttagagtct	23820
aatttgacac	cttggaanaac	ctcggagtat	tttctcgcgc	caactttatt	ttcaatttca	23880
tagtcccttg	tatttgtgga	ccaatcaatt	acagtatttt	cataatccca	atattcctcg	23940
gttctttgtt	tattgatatt	cgtataaaca	cgagcctctg	accataccct	gcatttcata	24000
gttcgaatca	aatattcctt	ctacaacccc	ctatttttgt	ctgtatatatt	gacttcttaa	24060
gtccctattc	agaaaaactcc	tagtaggttaa	ttgcaataca	ccgagtatca	acagtatatg	24120
aagtccttta	atccggttta	ctacaggact	gcgaggttcg	tctaattgct	actgaaattt	24180
tccaagcggt	gggcatattt	ttgatttttc	gctttttctt	cagtggtcgt	tccgagtgat	24240
ttttgatagc	gctttttctac	cacttagtgc	tgctgtatgt	atttatatga	tataccagcg	24300
atacatgtaa	atatttctat	aaagactgta	aacctctgat	tacctctgcc	tgtgcgtctt	24360
tattcgcttc	ctcagcagag	gaagcaatgg	ctgcgttctt	tgtttggtgg	aagatatctc	24420
ttgtcttttc	gaaataaaca	gtttccaata	gattttctcat	ctgagactca	atgtcctcaa	24480
ttaaggaaacc	caaattagcg	acatgtgatg	tgaaaataac	atctaattga	cgggacatat	24540
caatagcgat	gtcctttttcg	gtctgtcttg	tcagggtccc	ggataacatc	atatgagaat	24600
tctggtcagt	ttttgttttg	tccaagtgca	gaatgattgt	agtgggtgact	ctatagttga	24660
aactgtcggg	agaagaagga	gatgttgtaa	cttcaaaaac	atggatactg	tccaatttac	24720
tgtgatctga	ttgggtttttt	ttgaaaagca	ctacccctgc	aaaatcgtgc	ccattgaaat	24780
cttcttcatt	gaggtcccag	aggtacacgc	tgagataacc	gccttcataa	tagagatctc	24840
tgtaaacgtc	gaaagagtca	ttggctaaga	tctccaattt	tcttaaaggg	gctgagggaa	24900
aggggctgtc	ttgtagatcc	tttggggata	gttctgggta	ataagtgttc	gaccaaggcg	24960
atctgaacga	atcaatatca	cgattatagt	cgcagcataa	gtactcccgg	tttgaatcgg	25020
cggaatcttt	ctgtgtggat	aggggaacgt	ctactgaaga	tagtaaatct	tgtgccaatc	25080
ttgggtgttaa	ttcgatcaag	ttgttttagat	tctcctgtaa	cgtggtagga	tttagcctcc	25140
taagaagatc	taaagcagca	tcgaattgag	catcagacat	tacgggtggg	attgtagagt	25200
tggagagagt	ttttgtatgt	cacttgattg	aacagttctc	ctttatttgt	tacaaatgct	25260
ctcatTTTTa	cgttttccgt	ggtgaaaaaa	aaaaaaaaaa	atgaacatat	atacatgtag	25320
taacagcagt	agtagatgaa	gaatttagag	tttaaacagg	taacataacg	ctataaggga	25380
aaggaaaagt	gagatgttat	acatttatatt	tacgtagacg	ttgaagggaag	aagaagaaa	25440
aagggaacag	aagaagaaga	aaaagagaaa	ggaaattcat	gtggatttaa	gatcgcttcc	25500
cctttttact	taatgtcttg	taggatcatt	gagctttaat	acgtctactg	caggacccag	25560
taaacggtga	aaaccacttt	tccccaatgt	ggcaactttg	gttctctttg	tagcagtcac	25620
agtggcctgt	ctgggcaaat	cgtttagcaa	ggcacttca	ccgaaataat	catgggtctt	25680
cagtttattt	atgacacctt	ggcccttctt	agagacgtcc	acagctccgt	actcaattaa	25740
ataaaagtct	tccccttgat	caccctcgcg	aatgtattgt	tcaccgggct	ggtagatctt	25800
ggtatccagt	gcacgcgcaa	gtttggcacg	gtcgtacgta	gtcaaaactct	tcaaaactgg	25860
catgctcttc	aaaagatcgt	catacatgag	tctcttcttg	aaagagctgc	ccaaaagtat	25920
ttttctgaag	gtgagcctgt	ctagagccca	caacaaacag	tcggaggttg	ctacaacggg	25980

ggcagcacga	gggctgttgt	acataagagc	aagttccccg	aaactgggagc	ctggccccga	26040
agagttgacc	ttgtttgtcgt	tgacgtagaa	gtcaaacagta	cccttttcga	cgacatagaa	26100
gtagtccccct	tgggtcacctt	gcttgattat	cgtagcacct	ttggggagcg	acttctcctc	26160
cagacaattt	atgaccagcc	tttttgagtc	ggaatccagc	ttgttgaaca	gaaagtatt	26220
acggatcgat	ttttccagtc	tttgcaattg	ctgctcggac	ttttccttat	agtgatctgg	26280
agtccaatcg	tcaaaattgt	ttggttgtaa	ggctcaccca	ctaacagaag	tacgcctttg	26340
ggcgttgaag	tgcattggga	gtggaggagt	tgatgtcttt	tctctagtat	gctgttgttc	26400
ttcctgtgct	tgctggtgag	tgctcctgttc	gtgcgggtct	aaattaaacc	cactttttaa	26460
cacgttggag	tgtgggtcct	cgttcacaaa	gggggatttg	aacataacac	tcgatctgga	26520
tcttgattga	gattgagctg	attgaggtct	ggaaaatgac	tcctctggtt	ccggaaatag	26580
aacaatgttc	tttgccctaa	attcaggctc	cctggccttg	aggaacgctc	tctgttgttc	26640
cagcctttta	ttgaaatagt	tggcggagaa	ctgacgaaag	tcggacggat	tagcggcgtt	26700
gatttcgttc	tggaaacagt	gcaattcggc	ttgcgattcc	ttgggcaaag	aagataccat	26760
cgttttattct	tactgtttgc	ttttgaaaat	aatctgcttg	ttgtaaatga	tcttcctatt	26820
tatgtatgcy	tgaaatgcgt	gtaatggatg	tgatgatagc	gatcgcttga	cttagtcgag	26880
gaaagcgtaa	agtgtccctt	tttctctttc	ccttttccct	gtttgttcaa	atthtctctt	26940
ttctaccttc	ctcttcgttt	tgctccttgc	cagcaaaacg	aggaaggaca	agaattttgg	27000
agagcggcgg	cgcggcgcca	taataaatgc	acttgagcaa	atgcaataag	ttcagcttat	27060
ggcttctgca	acaaagaaga	ggcgggaatc	cgagcacagt	ataatcaata	tattatcagg	27120
cagaacaaaa	gtggcgggta	taacacgaag	aaagaaggaa	gggtgggtgat	tgggtgcttt	27180
gtttttttcc	gctcgtcgtg	gtggcatgcy	cttttttttt	cttgggtgaa	gattagccgc	27240
cgaagtcgta	tgctcctctt	ttgagcgttt	tgaagatag	aaaaaagctg	agtaagtaaa	27300
atcgtcggcg	gctaactctga	aaagtatata	aagtccttta	tattttccac	tcgtattctt	27360
ttcctccttc	cttttttttt	ttactttctc	acgtatggct	tgtttcattt	attatcagat	27420
cagtaatcac	gtgccgtgtc	gttagaaatg	acatttggtt	ctttgccttc	tttgtccttt	27480
ttctatcttc	ctactaagag	aaactattat	acaatacaat	aagagaacat	cgtgaaacaa	27540
cataaaaagat	atcacgtcat	cagagagaga	taaaacatca	caaccaatta	cattacagtg	27600
tattcgtttc	ttcttagaga	gagaaataaa	gagcgggata	aaaggacact	acacagtata	27660
ccccatcccc	aggtgcaata	atacacgtat	atctatgttt	attgcacatc	aaaccccaca	27720
tatatgtctg	ccagaaaacg	caagtttaat	agtctcaaac	cgctagacac	cttgaacagc	27780
tctcgtgcc	gctctccaag	gtcctctgct	tctctaccgc	ccaaaagata	taacactttt	27840
cgtaaggatc	cgaaaatcgt	tgatcatcta	aacaatgctt	ctacaaagga	tttcttacca	27900
gttttgagta	tgaacagcga	gagtaagagg	cagatcgaat	tgtcggataa	tgatgttgat	27960
aataatgatg	aaggtgaagg	cgtcaacctg	ggctgttcag	atcaagattt	tgaacctttg	28020
caaagctctc	ctttgaaaag	acactcatca	ctcaaaagca	cttccaatgg	tcttttgttc	28080
caaagtgtcta	ataatctggg	gaatggttca	ccggaaccgg	cagtagcgag	cacttctcca	28140
aatggctcaa	ttatttccac	taaactaaat	ttgaacggcc	aattttcttg	cgttgattcg	28200
aaaacattgc	gaatttatcg	gcataaagca	ccatgcataa	tgacttttgt	ctcagatcat	28260
aatcattccga	aattttcatt	gtattttcaa	caatcggtga	tctacaattc	acaagttaat	28320
ctgcttgatg	atggtgaatt	gataatttta	gataagaaga	actcttttat	ggctataatt	28380
ttaaaagatc	tgaaaaaggt	caagatgata	ctagacgtga	ataactcttc	aatcaacatt	28440
aacacgaaca	tcttgatatg	gtccactgca	agctccgctt	caaataaaaa	aataaagtct	28500
attaaaagat	tcctgttgat	gtcatattct	tcgtcgataa	aagtcgaaat	tttagatcat	28560
aaagagcaga	ttttggaaag	actaaaacat	ctgattcatc	ctatttcttc	gtcatcacct	28620
tcattgaaca	tggaaagggc	aataaactcc	actaaaaatg	cattcgactc	tttaagactt	28680
aaaaaaacta	aactttctac	taatgatgat	gaaagtccgc	aaattcatac	tcatttctta	28740
tcaaacaaac	ctcatggttt	gcaatcctta	acaaaaagga	ctcgtattgc	cagccttggg	28800
aaaaaagagc	attcaatatc	tgttccaaaa	tcgaatat	caccctcaga	tttctacaac	28860
actaacggga	cagaaaacttt	acaatcccac	gcagtttcac	aactaagacg	ttcaaataga	28920
tttaaagatg	tttcggatcc	agcaaaactca	aattcaaatt	cagaatttga	tgatgcaact	28980
acggaatttg	aaacaccaga	actgttttaa	ctagcctct	gttacaattt	caacgatggg	29040
tcaagctata	ctataacaaa	tcaagatttc	aagtgtcttt	tcaataagga	ttgggttaatt	29100
gatagcattt	tggatttttt	tacaaaattt	tacattgaat	catctattga	aaagtcaatt	29160
atcaaaaagag	agcaagttca	cttaatgtcc	tctttttttt	acacgaaact	aattagtaat	29220
ccagcagatt	attattctaa	tgtaaaaaag	tgggttaaca	atactgattt	gttttctaaa	29280
aagtacgttg	ttataccaat	taacataagt	tatcattggg	ttagttgcat	tataacaaac	29340
ttggatgcga	tcttgatttt	tcatcaaaa	aaagataaaa	acgatgccat	caactccgat	29400
gagatttcta	tcaataaatcc	tctgggtta	attttgactt	tcgactcgtt	gaggcaaac	29460
cattcaagag	aaattgatcc	aataaaaagaa	tttctcatat	cctatgcact	tgataaatat	29520
tcaattcaac	tggataaaa	acaaatcaaa	atgaaaacgt	gtccagttcc	acaacaacct	29580

aatatgagcg	attgtggtgt	tcatgttatt	ttgaatatta	gaaaatTTTT	tgaaaatccg	29640
gtggaaacaa	ttgatgtatg	gaagaattct	aagattaaaa	gtaagcactt	caccgcaaaa	29700
atgattaata	aatatTTTTga	taaaaatgaa	agaaatagtg	cgagaaagaa	tttaaggcat	29760
actctaaaac	tcttacaact	caattacatc	agctatctga	aaaaagaaaa	tttatatgaa	29820
gaagtatatgc	aaatggagga	gaaaaaaaagc	accaatatca	acaataatga	aaattacgac	29880
gatgacgatg	aagaaattca	aatcattgaa	aatatagatc	aaagcagtaa	agataataac	29940
gcgcagttaa	cctcggaacc	tccctgctca	cgatcatcca	gtatttcaac	aacagaaaga	30000
gagccgacag	agttgcataa	ttccgtagta	cgacagccca	ccggtgaaat	aataactgat	30060
aatgaagatc	ccgttcgcgc	tgcttctcca	gaaacagcat	ccgtttctcc	tcccatacgt	30120
cacaatatTT	taaagagttc	atctccttcc	atatcagaaa	gtgcgaatga	aactgaacaa	30180
gaagaatttta	cgtcaccata	ctttggaagg	ccgtcttttaa	agacgagggc	taagcagtTT	30240
gaaggcgTTT	cgtcaccaat	aaaaaacgat	caaggccctgt	catccacca	tgatattatg	30300
atgccgtcgc	ctaaaccta	aaggatttac	ccaagcaaaa	aaatcccaca	actttcctcg	30360
catgttcaat	ctctatcgac	tgattctatg	gaacgccaat	ctagtcccaa	taataccaac	30420
attgtaattt	cagatacaga	acaagattcg	aggttaggag	tgaattctga	aagtaaaaa	30480
actagtggta	ttgttaacag	ggatgactcg	gacgtcaatt	taattggtag	ttcactacct	30540
aatgtagcgg	aaaaaatca	tgacaacact	caggaaagta	atggtaataa	tgacagtctt	30600
ggtaaaatac	ttcaaaatgt	tgacaaagaa	ctgaacgaaa	agttgggtga	tattgatgat	30660
gtggcatttta	gtagtccaac	taggggtatt	ccaagaacca	gcgcgacaag	caaggggatca	30720
aatgcacaac	ttctctctaa	ttatggagat	gaaaataacc	agtctcaaga	ttctgtttgg	30780
gatgagggca	gagataatcc	tatactcttg	gaagatgaag	acccttgaaa	gcactcgcac	30840
gcatagtttt	atatTTTTTT	ttcttttctt	ctttatTTTT	atttttttaa	cattctcgca	30900
taataatagt	aacagcaata	aaataactag	taatatacat	actcacataa	ataggcatat	30960
tttttatTTT	gttctttcgt	attcagaatg	gtaaatTTTT	ctaatagaat	gttttcaaaa	31020
aattttTTTT	ctataattct	aacaataaga	aataaaacag	tggaaaaaca	tataacttaa	31080
tgtagatata	tatatgtaaa	tatgctagca	ttcatttttaa	atgtaaggaa	gaaaacgcct	31140
ttaactttca	tctggtaaat	tttctaaagc	tctaccctta	gtgtaaacct	cgtcttttac	31200
attctgtaca	gtacttttga	accattgaat	gaacagcaca	gttgcttgga	aaaagaacgt	31260
aaatgactct	aatgctatca	gtagtccgaa	ataaatatct	tgcatccttg	tccatggttg	31320
gaatgattcg	ttaagcgata	aagtgaact	cctggatcca	tggctaatat	cgataatctg	31380
ctttgtagaa	acatcacgta	aaaatactaa	ggataaaata	ataatcggca	agacgtaaat	31440
taccatgtgc	aggaaacctt	gtttaaacag	ttccaaagcc	attgttttca	agaagttcaa	31500
tttactagta	tttctgtttc	tgaaaaatag	tttcaatgac	ttatatccgg	ttccgaaatt	31560
gaccattgac	aaaacgacga	taacaaactc	tcgtcatatt	tcgtcatatt	taaacatttt	31620
ggaattacca	atagtgtatg	aagaattgcc	aaacactctg	atgggtgagat	aattgtaaca	31680
ggctagtata	ccatcgggtta	agacgaccca	cgaaaccata	gttctaataa	ccgcaatcat	31740
gatagctata	aacaccatca	ggttataact	caccaacata	ctatcaaata	tactgttaaa	31800
gatggccatg	aaaattcctt	gaccctcttc	cacattattg	atattattaa	cagggggcatt	31860
gttgatattc	ctctctggta	tatcgtgttc	cacttctctc	acttcgtttt	cgtcaacggc	31920
atgttcatcg	agagtgtttt	ggccttctgt	aaaatgaaga	atatccttag	caatcttga	31980
taaacaaaaca	gatccgacac	agacataata	aatgttcaaa	taggagtagc	taatgtcaga	32040
gagttccttc	catgcaacat	agaccttatt	tctttcacct	aacaatagct	ttacgactgg	32100
aagaaagccg	aaactacaaa	caaagtttat	taaagcttga	ctaataaatg	taacaccaag	32160
cattaaaatg	gaagcaaata	accaaaccag	gccaaagcaat	gtcatatagc	gtaaacggaa	32220
atctggaggg	acgtaaacga	tataatatct	atcatattcc	gtgttttgtt	cgtctaataa	32280
accaaattca	ccagcagcac	gtttgtttct	ctctttaatt	ctttccaaat	ctaagtgttt	32340
caacagctta	tcctcctttg	taacggggac	aaacattggt	tgtacataat	ttctggatac	32400
tatatcagaa	gaaggaacac	gcattaaaac	gccatcgggg	acaaaataag	catgaacgtc	32460
acgaacttgc	ccaaacaatt	cttctgcttg	ttctagagtc	tttgggttag	tgaaaagttc	32520
ttggtttgac	cattctgcat	tcttggcagc	aatgtacttg	taaaacaaat	tgcggtaaac	32580
aatatggcct	ctctctgttg	gggtatcttt	acctaagata	aaggatgata	aacgtaattt	32640
tcttgaacat	aacttaaaaa	ttgttttcca	atacctttcc	aataagggtt	taacgtaaga	32700
gctagactcc	aggatcctct	tcgtaaaata	taaagttagc	aaaattgtat	tgaacttcca	32760
tgaaataatt	gatgttggtt	tataagcttc	aggaacggat	aataagttcg	atttgagcat	32820
gaatgggaaa	aaaattctag	tatggaagcc	aaaacctaa	acaataaaaa	tcgcataaat	32880
aaacatggat	aagcaaagcc	ttgataattg	aatactcatc	ggatgaatta	aactatcatg	32940
caagatctta	atattaggat	cttctggaga	tctgataaag	aatagaacac	ctgggtcta	33000
gatatttttc	ctaatcattc	cgatatattt	ggcaaaccaa	tacatgtaga	gagtaacca	33060
cgtccagtat	acgaaaaggg	aaaatggagg	ccatattgctg	catattgaag	gaacccaaag	33120
catccgacta	ttggatgcca	ggattggggca	aaataatgag	aaatctagca	ttacaccagc	33180

```

tagaatgggg aatccagcca attcgatgaa aaataaagtg aacaccttga aagtgcattt 33240
gagtgcaaat aaaatctgga aaatcagcct tctagtaggg ttgctcattc cgttctcacg 33300
gccataaccc ctcgagacaa gattggatga agcacaaacg atactgactg aggttagata 33360
cgctcggttaa gcaggtaaag ctctgatgaa aattgagtgc ttcatgggtg ttccagtgtg 33420
gccgttgtaa agataaatga tgtcatgaat taagtgggtc gaaatccaac tcatcgcgac 33480
atctacttta gggaccagtt tcgtcagccc attgtaagct atatgagcgc cagacaagta 33540
gtaaagatgg cacaaccac gcaaaataac cttaaataa ccaaaataaa ttttcaacaa 33600
accgaaccca ataaaagttg gaaacaaata ggaaattgcc aggtatattg cggtgaaaac 33660
tacggcaatt atgaaataag caatgacatt caacaatttc aatttgagat ttatcactag 33720
agggccttgt ccttgatcgt ctgggttagc ctgtgctggt ggcacaccaa cagctgcgcc 33780
aaagtcttgt tcgtcttgat cgacatttcc agctctgttt tgggcagggt gtgggatgaa 33840
aaccggagcg ttgggcctgt ttattgcgtt ctgttgagca gcgattaggt catcaaattc 33900
attttgcgca cgacgatttg cccacatgtc aatcggattg ggctctatat gatc 33954

```

<210> 8

<211> 640

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 8

```

Met Arg Arg Glu Arg Gln Arg Met Met Ser Phe Glu Asp Lys Asp Lys
 1              5              10              15

Asp Asp Leu Asp Asn Ser Asn Ser Asn Asn Ser Ser Glu Met Thr Asp
      20              25              30

Thr Ala Met Met Pro Pro Leu Lys Arg Leu Leu Ile Thr Gly Ser Ser
      35              40              45

Asp Asp Leu Ala Gln Gly Ser Ser Gly Lys Lys Lys Met Thr Met Ala
      50              55              60

Thr Arg Ser Pro Ser Ser Ser Pro Asp Leu Ala Thr Asn Asp Ser Gly
      65              70              75              80

Thr Arg Val Gln Pro Leu Pro Glu Tyr Asn Phe Thr Lys Phe Cys Tyr
      85              90              95

Arg His Asn Pro Asp Ile Gln Phe Ser Pro Thr His Thr Ala Cys Tyr
      100              105              110

Lys Gln Asp Leu Lys Arg Thr Gln Glu Ile Asn Ala Asn Ile Ala Lys
      115              120              125

Leu Pro Leu Gln Glu Gln Ser Asp Ile His His Ile Ile Ser Lys Tyr
      130              135              140

Ser Asn Ser Asn Asp Lys Ile Arg Lys Leu Ile Leu Asp Gly Ile Leu
      145              150              155              160

Ser Thr Ser Cys Phe Pro Gln Leu Ser Tyr Ile Ser Ser Leu Val Thr
      165              170              175

His Met Ile Lys Ile Asp Phe Ile Ser Ile Leu Pro Gln Glu Leu Ser
      180              185              190

Leu Lys Ile Leu Ser Tyr Leu Asp Cys Gln Ser Leu Cys Asn Ala Thr
      195              200              205

```

Arg Val Cys Arg Lys Trp Gln Lys Leu Ala Asp Asp Asp Arg Val Trp
 210 215 220
 Tyr His Met Cys Glu Gln His Ile Asp Arg Lys Cys Pro Asn Cys Gly
 225 230 235 240
 Trp Gly Leu Pro Leu Leu His Met Lys Arg Ala Arg Ile Gln Gln Asn
 245 250 255
 Ser Thr Gly Ser Ser Ser Asn Ala Asp Ile Gln Thr Gln Thr Thr Arg
 260 265 270
 Pro Trp Lys Val Ile Tyr Arg Glu Arg Phe Lys Val Glu Ser Asn Trp
 275 280 285
 Arg Lys Gly His Cys Arg Ile Gln Glu Phe Lys Gly His Met Asp Gly
 290 295 300
 Val Leu Thr Leu Gln Phe Asn Tyr Arg Leu Leu Phe Thr Gly Ser Tyr
 305 310 315 320
 Asp Ser Thr Ile Gly Ile Trp Asp Leu Phe Thr Gly Lys Leu Ile Arg
 325 330 335
 Arg Leu Ser Gly His Ser Asp Gly Val Lys Thr Leu Tyr Phe Asp Asp
 340 345 350
 Arg Lys Leu Ile Thr Gly Ser Leu Asp Lys Thr Ile Arg Val Trp Asn
 355 360 365
 Tyr Ile Thr Gly Glu Cys Ile Ser Thr Tyr Arg Gly His Ser Asp Ser
 370 375 380
 Val Leu Ser Val Asp Ser Tyr Gln Lys Val Ile Val Ser Gly Ser Ala
 385 390 395 400
 Asp Lys Thr Val Lys Val Trp His Val Glu Ser Arg Thr Cys Tyr Thr
 405 410 415
 Leu Arg Gly His Thr Glu Trp Val Asn Cys Val Lys Leu His Pro Lys
 420 425 430
 Ser Phe Ser Cys Phe Ser Cys Ser Asp Asp Thr Thr Ile Arg Met Trp
 435 440 445
 Asp Ile Arg Thr Asn Ser Cys Leu Lys Val Phe Arg Gly His Val Gly
 450 455 460
 Gln Val Gln Lys Ile Ile Pro Leu Thr Ile Lys Asp Val Glu Asn Leu
 465 470 475 480
 Ala Thr Asp Asn Thr Ser Asp Gly Ser Ser Pro Gln Asp Asp Pro Thr
 485 490 495
 Met Thr Asp Gly Ala Asp Glu Ser Asp Thr Pro Ser Asn Glu Gln Glu
 500 505 510
 Thr Val Leu Asp Glu Asn Ile Pro Tyr Pro Thr His Leu Leu Ser Cys
 515 520 525

Gly Leu Asp Asn Thr Ile Lys Leu Trp Asp Val Lys Thr Gly Lys Cys
 530 535 540
 Ile Arg Thr Gln Phe Gly His Val Glu Gly Val Trp Asp Ile Ala Ala
 545 550 555 560
 Asp Asn Phe Arg Ile Ile Ser Gly Ser His Asp Gly Ser Ile Lys Val
 565 570 575
 Trp Asp Leu Gln Ser Gly Lys Cys Met His Thr Phe Asn Gly Arg Arg
 580 585 590
 Leu Gln Arg Glu Thr Gln His Thr Gln Thr Gln Ser Leu Gly Asp Lys
 595 600 605
 Val Ala Pro Ile Ala Cys Val Cys Ile Gly Asp Ser Glu Cys Phe Ser
 610 615 620
 Gly Asp Glu Phe Gly Cys Val Lys Met Tyr Lys Phe Asp Leu Asn Asp
 625 630 635 640

<210> 9
 <211> 2112
 <212> DNA
 <213> *Saccharomyces cerevisiae*

<400> 9
 atgtcactct ctaggtgtcc aactgacaat tcgtcctccc gtataaattc ctctgttccct 60
 ttaataaata gcagcagccc tgctacacct ccagaatcat ttgatcctca agtatttcct 120
 tcttcactta ttcattggga taacctgctt cctcaagatg atcaaattgc atcggtacct 180
 cgctcagaat caaatagttg taatggcaat acgagttctt ccctgcccgtg cactgattcg 240
 tatcagtacc cattaagca ttcttgtagc ccttcttttc ttcgaaagt taatgaaagt 300
 atagagaatg tctcttataa atgcttagac cactcacgc cagatagtgt tcctggcgat 360
 ttttccattt cccttgttcc tcaaaggaat tttctatatt ctcatctctc tcttccacct 420
 aaaattatat caattgatag aaacaatcga attaatgttag ataatagcat ttcacttaac 480
 tccgacaatt tccctccttc tccgaaagtc gacacatcaa acactgtttc acctggtagt 540
 aaacctatct ctgaggatct tgaagattta aacttacagt caattgttca aacttttgag 600
 gatcttccag aaggaattca atcttatgag ttttttcaac tactccgttc gtgcaatcgg 660
 caatcgatgc gtttattatt gaattgaatgc gagccgcttc taaaaaaaga tatactttca 720
 aatcttccct tttccattgt tcagtctata ttattaaatc tggatataca tttttttctt 780
 tcttgccgtc ttgtttcgcc tacttggaat agaatacttg atgtgcatac ttcatactgg 840
 aaacacatgt ttagtttatt tggctttcaa atcaatgaaa atgactggaa atatgctaatt 900
 ccaaacttaa atcgccacc ttttttgcac aacgaccaa tctcagatga ctattttccg 960
 gaaattttca aaagacattt tctcaataga aaacgatggt tatttccttc gatactcca 1020
 agtcatctat cttttcccat tcatgttcca aactttatga taacttcttt actacttcat 1080
 aaagacagaa taatcaccac ttcgggatct ggaacaattc aaattcataa tgctattacc 1140
 ggtgttttag aagctcgatt agagggtcat aaagaagggtg tttgggctgt caaaatacat 1200
 gagaatacac ttgtatctgg ttcgatcgat aaaactgttc gcgtttggaa catagagaaa 1260
 gctaaatgta cgcacatatt taggggacat atttccatca tcagatgctt agagatctta 1320
 gtcccgagtc gtcttattcg ccatggagtt gaaattgttg aaccagatca accgtacatt 1380
 gtccgaggct tcgggatca tacacttcgg gtttgggaagc ttccaaaaaa cacggtacct 1440
 ccttatcttc cagataatac aaactctatt gaccgttggg agaagaaccg gtattttgta 1500
 catactttga taggacatac agactctgta cgaactatat ccggctatgg tgatatactt 1560
 gtgagtggga gttatgattc ttcaattcgc atttggagag tttcaacagg ggaatgtctt 1620
 taccatctgc ggggtcatag tcttcgtata tatagtgttt tatatgaacc agaaaggaat 1680
 atttgcataa gcggtagcat ggataagtc attagggttt gggatttatc gacagggaact 1740
 tgtaaatatg tgcttgaagg ccatgatgcc tttgttacgc ttcttaattg attccagaat 1800
 aggttgatat caggttctgc tgactccaca attagaatat gggatttgaa tactgggaaa 1860
 ccattaatgg ttttgccgtc taattcaggc tacattagta gctttgtgtc agatgaacac 1920

```

aaaattatta gtggtaatga tgggttctgta aagttatggg atggttaggac tggaaagctg 1980
ttacgttttc tattaacaga cctcacaaaa atatggcatg tgcattttga tgctatgcgt 2040
tgtgtggctg cagtgcagcg tgatgatcaa gcatatttgg aagttattaa tttttccgga 2100
tcaagaccgt ag 2112

```

<210> 10

<211> 703

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 10

```

Met Ser Leu Ser Arg Cys Pro Thr Asp Asn Ser Ser Ser Arg Ile Asn
 1              5              10              15

Ser Ser Val Pro Leu Ile Asn Ser Ser Ser Pro Ala Thr Pro Pro Glu
      20              25              30

Ser Phe Asp Pro Gln Val Phe Pro Ser Ser Leu Ile His Gly Asp Asn
      35              40              45

Leu Leu Pro Gln Asp Asp Gln Ile Ala Ser Asp Pro Arg Ser Glu Ser
 50              55              60

Asn Ser Cys Asn Gly Asn Thr Ser Ser Ser Leu Pro Cys Thr Asp Ser
 65              70              75              80

Tyr Gln Tyr Pro Leu Lys His Ser Cys Thr Pro Ser Phe Leu Arg Lys
      85              90              95

Phe Asn Glu Ser Ile Glu Asn Val Ser Tyr Lys Cys Leu Asp His Ser
      100              105              110

Pro Pro Asp Ser Val Pro Gly Asp Phe Ser Ile Ser Leu Val Pro Gln
      115              120              125

Arg Asn Phe Leu Tyr Ser His Ser Ser Leu Pro Pro Lys Ile Ile Ser
      130              135              140

Ile Asp Arg Asn Asn Arg Ile Lys Leu Asp Asn Ser Ile Ser Ser Asn
      145              150              155              160

Ser Asp Asn Phe Pro Pro Ser Pro Lys Val Asp Thr Ser Asn Thr Val
      165              170              175

Ser Pro Gly Ser Lys Pro Ile Ser Glu Asp Leu Glu Asp Leu Asn Leu
      180              185              190

Gln Ser Ile Val Gln Thr Phe Glu Asp Leu Pro Glu Gly Ile Gln Ser
      195              200              205

Tyr Ala Phe Phe Gln Leu Leu Arg Ser Cys Asn Arg Gln Ser Met Arg
      210              215              220

Leu Leu Leu Asn Glu Cys Glu Pro Leu Leu Lys Lys Asp Ile Leu Ser
      225              230              235              240

Asn Leu Pro Phe Ser Ile Val Gln Ser Ile Leu Leu Asn Leu Asp Ile
      245              250              255

```

His Ser Phe Leu Ser Cys Arg Leu Val Ser Pro Thr Trp Asn Arg Ile
 260 265 270
 Leu Asp Val His Thr Ser Tyr Trp Lys His Met Phe Ser Leu Phe Gly
 275 280 285
 Phe Gln Ile Asn Glu Asn Asp Trp Lys Tyr Ala Asn Pro Asn Leu Asn
 290 295 300
 Arg Pro Pro Phe Leu His Asn Asp Gln Ile Ser Asp Asp Tyr Phe Pro
 305 310 315 320
 Glu Ile Phe Lys Arg His Phe Leu Asn Arg Lys Arg Trp Leu Phe Pro
 325 330 335
 Ser Ile Pro Pro Ser His Leu Ser Phe Pro Ile His Val Pro Asn Phe
 340 345 350
 Met Ile Thr Ser Leu Leu Leu His Lys Asp Arg Ile Ile Thr Thr Ser
 355 360 365
 Gly Ser Gly Thr Ile Gln Ile His Asn Ala Ile Thr Gly Val Leu Glu
 370 375 380
 Ala Arg Leu Glu Gly His Lys Glu Gly Val Trp Ala Val Lys Ile His
 385 390 395 400
 Glu Asn Thr Leu Val Ser Gly Ser Ile Asp Lys Thr Val Arg Val Trp
 405 410 415
 Asn Ile Glu Lys Ala Lys Cys Thr His Ile Phe Arg Gly His Ile Ser
 420 425 430
 Ile Ile Arg Cys Leu Glu Ile Leu Val Pro Ser Arg Leu Ile Arg His
 435 440 445
 Gly Val Glu Ile Val Glu Pro Asp Gln Pro Tyr Ile Val Ser Gly Ser
 450 455 460
 Arg Asp His Thr Leu Arg Val Trp Lys Leu Pro Lys Asn Thr Asp Pro
 465 470 475 480
 Pro Tyr Leu Pro Asp Asn Thr Asn Ser Ile Asp Arg Trp Glu Lys Asn
 485 490 495
 Pro Tyr Phe Val His Thr Leu Ile Gly His Thr Asp Ser Val Arg Thr
 500 505 510
 Ile Ser Gly Tyr Gly Asp Ile Leu Val Ser Gly Ser Tyr Asp Ser Ser
 515 520 525
 Ile Arg Ile Trp Arg Val Ser Thr Gly Glu Cys Leu Tyr His Leu Arg
 530 535 540
 Gly His Ser Leu Arg Ile Tyr Ser Val Leu Tyr Glu Pro Glu Arg Asn
 545 550 555 560
 Ile Cys Ile Ser Gly Ser Met Asp Lys Ser Ile Arg Val Trp Asp Leu
 565 570 575

Ser Thr Gly Thr Cys Lys Tyr Val Leu Glu Gly His Asp Ala Phe Val
580 585 590

Thr Leu Leu Asn Val Phe Gln Asn Arg Leu Ile Ser Gly Ser Ala Asp
595 600 605

Ser Thr Ile Arg Ile Trp Asp Leu Asn Thr Gly Lys Pro Leu Met Val
610 615 620

Leu Pro Ser Asn Ser Gly Tyr Ile Ser Ser Phe Val Ser Asp Glu His
625 630 635 640

Lys Ile Ile Ser Gly Asn Asp Gly Ser Val Lys Leu Trp Asp Val Arg
645 650 655

Thr Gly Lys Leu Leu Arg Phe Leu Leu Thr Asp Leu Thr Lys Ile Trp
660 665 670

His Val Asp Phe Asp Ala Met Arg Cys Val Ala Ala Val Gln Arg Asp
675 680 685

Asp Gln Ala Tyr Leu Glu Val Ile Asn Phe Ser Gly Ser Arg Pro
690 695 700

<210> 11
<211> 2175
<212> DNA
<213> Mus sp.

<400> 11
gaattcggca cgaggcggag ctgcgttggc tgcggcctgg cagcaaaggg gcggccccgg 60
cggagagcag acccagtagt ccgggcgatt atggaccggg cagaggcggg gctgcaggag 120
aaagcgctta agtttatgaa ttccctcagag agagaagact gtaataatgg cgaaccccct 180
aggaagataa taccagagaa gaattcactt agacagactt acaacagctg tgccaggctt 240
tgcataaacc aagagacagt atgtctaaca agcactgcta tgaagactga aaattgtgtg 300
gccaaagcca aacttgccaa tggcacttcc agcatgattg tgcccaagca gcggaaactc 360
tcagcaagct atgagaagga aaaggagctg tgtgtcaagt attttgagca gtggtcagag 420
tctgatcaag tgggaatttgt agaacacctt atatcccaaa tgtgtcacta ccagcatggg 480
cacatcaact cctacctaaa acctatgctg cagaggggatt tcataactgc actgccagca 540
cgggggtctgg accacatcgc tgagaacatt ctgtcatact tggacgccaa gtactgtgt 600
gctgctgagc tcgtgtgcaa ggaatggtac cgcgtgacgt cggacggcat gctgtggaaa 660
aagctcatcg agaggatggg caggacggac tctctgtggc gaggcctggc agagcgaga 720
ggctgggggac agtacttatt caaaaacaaa cctcctgatg agaacgctcc tcccaactcc 780
ttttatagag cgctttatcc taaaatcata caagacattg agacaataga gtccaattgg 840
agatgtggggc gacatagttt acagagaatc cactgccgga gtgaaacaag taaagggggt 900
tactgtttac agtacgacga ccagaagata gtcagcggcc ttcgagacaa caccatcaag 960
atctgggata aaagcacact ggaatgcaag cggatttctca cgggccacac gggctccgtc 1020
ctgtgtctgc agtacgatga gaggtgatc atcacaggct cctcagactc caccgtcaga 1080
gtgtgggatg taaatgcagg tgagatgcta aacacattga ttcaccactg tgaagccgtt 1140
ctgcacctgc gcttcaataa tggcatgatg gtgacctgtt ccaaagaccg ttccatcgct 1200
gtgtgggata tggcttcccc aactgacatc accctcagga ggggtgctggg gggacaccga 1260
gctgcggtca atgtttaga ctttgatgac aagtacatcg tttctgcctc tggagataga 1320
accataaagg tctggaacac aagtacctgt gaattcgtaa ggaccctaaa tgggcacaac 1380
cgtggcatcg cctgtttgca gtacagagac aggcctgggtg tgagcggctc ctctgacaac 1440
accatcaggc tgtgggacat agagtgtgga gcatgcctgc gagtgttggg gggccatgag 1500
gagttggtac gctgcattcg atttgataac aaaaggatag tgagcggagc ctatgatggg 1560
aaaattaaag tgtgggatct tatggctgct ttggaccgcg gtgctccagc agggactctc 1620


```

tgtctgcgga cacttggtga gcattctgga agagttttcc gcctccagtt tgatgaattc 1680
cagattgtca gtagttcaca tgatgacaca attctcatct gggacttcct gaatgatcca 1740
gctgctcacg ctgaaccgcc ccgctcccct tctcggacat acacctacat ctccagataa 1800
ataacccaac actggcctca taattgccca ggattcggtta atgttgagc atttaacaga 1860
cctgccaaga ccaggatgaa caacaatcaa actcctaccc ggattcccgg acggatgagc 1920
gaggagcagg gctttgagac tcctgttggg acacagtcgg tcagcagccg accaggacgg 1980
cctgctcggc accggctgcc tcagtgtctgc tatcagaaga tgtctttatc ttgtgtgaat 2040
gattggaact tccaagcctc cctccccttc cctccccttc cctccctgca cctgtttccc 2100
tcccattggg ttccagacaa agatgactta taaatatatt tagtgttttg cctaaaaaaaa 2160
aaaaaaaaaa aaaaaa                                     2175

```

<210> 12
 <211> 569
 <212> PRT
 <213> Mus sp.

```

<400> 12
Met Asp Pro Ala Glu Ala Val Leu Gln Glu Lys Ala Leu Lys Phe Met
 1          5          10          15

Asn Ser Ser Glu Arg Glu Asp Cys Asn Asn Gly Glu Pro Pro Arg Lys
      20          25          30

Ile Ile Pro Glu Lys Asn Ser Leu Arg Gln Thr Tyr Asn Ser Cys Ala
      35          40          45

Arg Leu Cys Ile Asn Gln Glu Thr Val Cys Leu Thr Ser Thr Ala Met
      50          55          60

Lys Thr Glu Asn Cys Val Ala Lys Ala Lys Leu Ala Asn Gly Thr Ser
      65          70          75          80

Ser Met Ile Val Pro Lys Gln Arg Lys Leu Ser Ala Ser Tyr Glu Lys
      85          90          95

Glu Lys Glu Leu Cys Val Lys Tyr Phe Glu Gln Trp Ser Glu Ser Asp
      100          105          110

Gln Val Glu Phe Val Glu His Leu Ile Ser Gln Met Cys His Tyr Gln
      115          120          125

His Gly His Ile Asn Ser Tyr Leu Lys Pro Met Leu Gln Arg Asp Phe
      130          135          140

Ile Thr Ala Leu Pro Ala Arg Gly Leu Asp His Ile Ala Glu Asn Ile
      145          150          155          160

Leu Ser Tyr Leu Asp Ala Lys Ser Leu Cys Ala Ala Glu Leu Val Cys
      165          170          175

Lys Glu Trp Tyr Arg Val Thr Ser Asp Gly Met Leu Trp Lys Lys Leu
      180          185          190

Ile Glu Arg Met Val Arg Thr Asp Ser Leu Trp Arg Gly Leu Ala Glu
      195          200          205

Arg Arg Gly Trp Gly Gln Tyr Leu Phe Lys Asn Lys Pro Pro Asp Glu
      210          215          220

```

Asn Ala Pro Pro Asn Ser Phe Tyr Arg Ala Leu Tyr Pro Lys Ile Ile
 225 230 235 240
 Gln Asp Ile Glu Thr Ile Glu Ser Asn Trp Arg Cys Gly Arg His Ser
 245 250 255
 Leu Gln Arg Ile His Cys Arg Ser Glu Thr Ser Lys Gly Val Tyr Cys
 260 265 270
 Leu Gln Tyr Asp Asp Gln Lys Ile Val Ser Gly Leu Arg Asp Asn Thr
 275 280 285
 Ile Lys Ile Trp Asp Lys Ser Thr Leu Glu Cys Lys Arg Ile Leu Thr
 290 295 300
 Gly His Thr Gly Ser Val Leu Cys Leu Gln Tyr Asp Glu Arg Val Ile
 305 310 315 320
 Ile Thr Gly Ser Ser Asp Ser Thr Val Arg Val Trp Asp Val Asn Ala
 325 330 335
 Gly Glu Met Leu Asn Thr Leu Ile His His Cys Glu Ala Val Leu His
 340 345 350
 Leu Arg Phe Asn Asn Gly Met Met Val Thr Cys Ser Lys Asp Arg Ser
 355 360 365
 Ile Ala Val Trp Asp Met Ala Ser Pro Thr Asp Ile Thr Leu Arg Arg
 370 375 380
 Val Leu Val Gly His Arg Ala Ala Val Asn Val Val Asp Phe Asp Asp
 385 390 395 400
 Lys Tyr Ile Val Ser Ala Ser Gly Asp Arg Thr Ile Lys Val Trp Asn
 405 410 415
 Thr Ser Thr Cys Glu Phe Val Arg Thr Leu Asn Gly His Lys Arg Gly
 420 425 430
 Ile Ala Cys Leu Gln Tyr Arg Asp Arg Leu Val Val Ser Gly Ser Ser
 435 440 445
 Asp Asn Thr Ile Arg Leu Trp Asp Ile Glu Cys Gly Ala Cys Leu Arg
 450 455 460
 Val Leu Glu Gly His Glu Glu Leu Val Arg Cys Ile Arg Phe Asp Asn
 465 470 475 480
 Lys Arg Ile Val Ser Gly Ala Tyr Asp Gly Lys Ile Lys Val Trp Asp
 485 490 495
 Leu Met Ala Ala Leu Asp Pro Arg Ala Pro Ala Gly Thr Leu Cys Leu
 500 505 510
 Arg Thr Leu Val Glu His Ser Gly Arg Val Phe Arg Leu Gln Phe Asp
 515 520 525
 Glu Phe Gln Ile Val Ser Ser Ser His Asp Asp Thr Ile Leu Ile Trp
 530 535 540

Asp Phe Leu Asn Asp Pro Ala Ala His Ala Glu Pro Pro Arg Ser Pro
 545 550 555 560

Ser Arg Thr Tyr Thr Tyr Ile Ser Arg
 565

<210> 13

<211> 9

<212> PRT

<213> Unknown Organism

<220>

<223> Description of Unknown Organism: SV40 Large T antigen fragment

<400> 13

Pro Pro Lys Lys Lys Arg Lys Val Ala
 1 5

<210> 14

<211> 32

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: formula
 sequence

<220>

<221> MOD_RES

<222> (3)...(6)

<223> any amino acid

<220>

<221> MOD_RES

<222> (7)

<223> hydrophobic amino acid

<220>

<221> MOD_RES

<222> (8)...(9)

<223> any amino acid

<220>

<221> MOD_RES

<222> (10)

<223> hydrophobic amino acid

<220>

<221> MOD_RES

<222> (11)

<223> any amino acid

<220>

<221> MOD_RES

<222> (12)

<223> aromatic amino acid

peptide

<400> 15

Cys Met His Ile Glu Ser Leu Asp Ser Tyr Thr Cys
 1 5 10

<210> 16

<211> 12

<212> PRT

<213> Unknown Organism

<220>

<223> Description of Unknown Organism: EGF-derived
 peptide

<400> 16

Cys Met Tyr Ile Glu Ala Leu Asp Lys Tyr Ala Cys
 1 5 10

<210> 17

<211> 32

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: formula
 sequence

<220>

<221> MOD_RES

<222> (1)

<223> unique amino acid, such as cys or lys

<220>

<221> MOD_RES

<222> (2)...(3)

<223> amino acid residue selected to modulate the affinity of
 the internalizing peptide for different membranes

<400> 17

Xaa Xaa Xaa Glu Ala Ala Leu Ala Glu Ala Leu Ala Glu Ala Leu Ala
 1 5 10 15

Glu Ala Leu Ala Glu Ala Leu Ala Glu Ala Leu Glu Ala Leu Ala Ala
 20 25 30

<210> 18

<211> 8

<212> PRT

<213> Unknown Organism

<220>

<223> Description of Unknown Organism: peptide substrate
 for N-myristoyl transferase

<400> 18

Gly Asn Ala Ala Ala Arg Arg
1 5

<210> 19

<211> 10

<212> PRT

<213> Unknown Organism

<220>

<223> Description of Unknown Organism: peptide derived
from laminin

<400> 19

Cys Asp Pro Gly Tyr Ile Gly Ser Arg Cys
1 5 10

<210> 20

<211> 75

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Nde1-EcoR1 fragment

<400> 20

catatgggtg gctgccgtgg cgatatgttc gggtgcggtg ctctccaaa aaagaagaga 60
aaggtagctg gattc 75

<210> 21

<211> 24

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: RGD/SV40 peptide

<400> 21

Met Gly Gly Cys Arg Gly Asp Met Phe Gly Cys Gly Ala Pro Pro Lys
1 5 10 15

Lys Lys Arg Lys Val Ala Gly Phe
20

<210> 22

<211> 225

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Nde1-EcoR1 fragment

<400> 22

catatgggagc cagtagatcc tagactagag ccctggaagc atccaggaag tcagcctaaa 60
actgcttgta ccaattgcta ttgtaaaaag tgttgctttc attgccaagt ttgtttcata 120

acaaaagccc ttggcatctc ctatggcagg aagaagcgga gacagcgacg aagacctoct 180
 caaggcagtc agactcatca agtttctcta agtaagcaag gattc 225

<210> 23
 <211> 72
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: engineered HIV-1 tat

<400> 23
 Met Glu Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser
 1 5 10 15
 Gln Pro Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe
 20 25 30
 His Cys Gln Val Cys Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly
 35 40 45
 Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro Gln Gly Ser Gln Thr
 50 55 60
 His Gln Val Ser Leu Ser Lys Gln
 65 70

<210> 24
 <211> 912
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: NdeI-EcoRI fragment

<400> 24
 catatgacct ctgcgcgtc cgtgaagtcg ggtccgcggg aggttccgcg cgatgagtac 60
 gaggatctgt actacacccc gtcttcaggt atggcgagtc ccgatatgcc gcctgacacc 120
 tccgcgcgtg ggcacctaca gacacgctcg cgccagaggg gcgaggtccg ttctgtccag 180
 tacgacgagt cggattatgc cctctacggg ggctcgtcat ccgaagacga cgaacacccg 240
 gaggtccccc ggacgcggcg tcccgtttcc ggggcgggtt tgtccggccc ggggcctgcg 300
 cgggcgcctc cgccaccgc tgggtccgga ggggcgggac gcacaccac caccgcccc 360
 cgggcccccc gaaccacgc ggtggcgact aaggcccccg cgccccggc ggcggagacc 420
 accgcgggca ggaaatcggc ccagccagaa tccgcgcac tccagacgc cccgcgctcg 480
 acggcgccaa ccgatccaa gacaccgcg caggggctgg ccagaaagct gcactttagc 540
 accgcccccc caaacccga cgcgccatgg accccccggg tggccggctt taacaagcgc 600
 gtcttctgcg ccgcggtcgg gcgcctggcg gccatgcatg ccgggatggc ggcgggtccag 660
 ctctgggaca tgtcgcgtcc gcgcacagac gaagacctca acgaactcct tggcatcacc 720
 accatccgcg tgacggtctg cgagggcaaa aacctgcttc agcgcgccaa cgagttgggtg 780
 aatccagacg tgggtgcagga cgtcgacgcg gccacggcga ctcgagggcg ttctgcggcg 840
 tcgcgcccc aagagcgacc tcgagcccca gcccgctccg cttctcgccc cagacggccc 900
 gtcgaggaat tc 912

<210> 25
 <211> 301
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: engineered HSV-1 VP22

<400> 25

Met Thr Ser Arg Arg Ser Val Lys Ser Gly Pro Arg Glu Val Pro Arg
 1 5 10 15

Asp Glu Tyr Glu Asp Leu Tyr Tyr Thr Pro Ser Ser Gly Met Ala Ser
 20 25 30

Pro Asp Ser Pro Pro Asp Thr Ser Arg Arg Gly Ala Leu Gln Thr Arg
 35 40 45

Ser Arg Gln Arg Gly Glu Val Arg Phe Val Gln Tyr Asp Glu Ser Asp
 50 55 60

Tyr Ala Leu Tyr Gly Gly Ser Ser Ser Glu Asp Asp Glu His Pro Glu
 65 70 75 80

Val Pro Arg Thr Arg Arg Pro Val Ser Gly Ala Val Leu Ser Gly Pro
 85 90 95

Gly Pro Ala Arg Ala Pro Pro Pro Pro Ala Gly Ser Gly Gly Ala Gly
 100 105 110

Arg Thr Pro Thr Thr Ala Pro Arg Ala Pro Arg Thr Gly Arg Val Ala
 115 120 125

Thr Lys Ala Pro Ala Ala Pro Ala Ala Glu Thr Thr Arg Gly Arg Lys
 130 135 140

Ser Ala Gln Pro Glu Ser Ala Ala Leu Pro Asp Ala Pro Ala Ser Thr
 145 150 155 160

Ala Pro Thr Arg Ser Lys Thr Pro Ala Gln Gly Leu Ala Arg Lys Leu
 165 170 175

His Phe Ser Thr Ala Pro Pro Asn Pro Asp Ala Pro Trp Thr Pro Arg
 180 185 190

Val Ala Gly Phe Asn Lys Arg Val Phe Cys Ala Ala Val Gly Arg Leu
 195 200 205

Ala Ala Met His Ala Arg Met Ala Ala Val Gln Leu Trp Asp Met Ser
 210 215 220

Arg Pro Arg Thr Asp Glu Asp Leu Asn Glu Leu Leu Gly Ile Thr Thr
 225 230 235 240

Ile Arg Val Thr Val Cys Glu Gly Lys Asn Leu Leu Gln Arg Ala Asn
 245 250 255

Glu Leu Val Asn Pro Asp Val Val Gln Asp Val Asp Ala Ala Thr Ala
 260 265 270

Thr Arg Gly Arg Ser Ala Ala Ser Arg Pro Thr Glu Arg Pro Arg Ala
 275 280 285

Pro Ala Arg Ser Ala Ser Arg Pro Arg Arg Pro Val Glu
 290 295 300

<210> 26
 <211> 120
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Nde1-EcoR1 fragment

<400> 26
 catatggacg tgcacgcggc cacggcgact cgagggcggtt ctgcggcgctc gcgccccacc 60
 gagcgacctc gagccccagc ccgctccgct tctcgcccca gacggcccgt cgaggaattc 120

<210> 27
 <211> 37
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: VP22 (C-terminal domain)

<400> 27
 Met Asp Val Asp Ala Ala Thr Ala Thr Arg Gly Arg Ser Ala Ala Ser
 1 5 10 15
 Arg Pro Thr Glu Arg Pro Arg Ala Pro Ala Arg Ser Ala Ser Arg Pro
 20 25 30
 Arg Arg Pro Val Glu
 35

<210> 28
 <211> 16
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Linker peptide

<400> 28
 Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser
 1 5 10 15

<210> 29
 <211> 4
 <212> PRT
 <213> Unknown Organism

<220>
 <223> Description of Unknown Organism: tetrapeptide isostere

<400> 29
 Ala Ile Tyr Tyr
 1

<210> 30
 <211> 38
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 30
 gcggatccac catggataam aaagagggac ctaataac 38

<210> 31
 <211> 76
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 31
 gcgcggccgc ctactcatca tctactagatg gcamcttctg agcaaaacag ccctctggta 60
 ttatagttgt cctcgt 76

<210> 32
 <211> 72
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 32
 cgcgggccgcc tactcatcat cactagatgg camttgagcc aaagttttct ctggtattat 60
 agttgtcctc gt 72

<210> 33
 <211> 39
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> modified_base
 <222> (20)
 <223> i

<220>
 <223> Description of Artificial Sequence: primer

<400> 33
 gcwatccacc atggataatn taaagaggga cctaataac 39

<210> 34
 <211> 34
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 34
 gtaggtgtat ctccatgtgg tatratagtr gtcc 34

<210> 35
 <211> 34
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 35
 ggacaactat aataccacat ggagatacac ctac 34

<210> 36
 <211> 29
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 36
 gcctcgagtc actcctcctc tgagctgtc 29

<210> 37
 <211> 38
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 37
 gcggatccac catggataam aaagaggac ctaataac 38

<210> 38
 <211> 37
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 38
 cctatcacat ctatatattta ttggtattat agttgtc 37

<210> 39
 <211> 37
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 39
 gacaactata ataccaataa aatatagatg tgatagg 37

<210> 40
 <211> 34
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 40
 gcctcgagtc ataatgtgtt agtatattgt cctg 34

<210> 41
 <211> 69
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 41
 gcggatccac catggaggaa gaagagtata tgcccatgga ggagactctt tgccaacggt 60
 ttaaattgtg 69

<210> 42
 <211> 35
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 42
 gcgcggccgc tcatatagac ataaatccag tagac 35

<210> 43
 <211> 65
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 43
 gcggatccgc caccatggac tacaaggacg acgatgacaa agatgaccgc gccgaggcgg 60

tgctg

65

<210> 44
 <211> 37
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 44
 gtaggtgtat ctccatgtct ggagatgtag gtgtatg

37

<210> 45
 <211> 37
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 45
 catacaccta catctccaga catggagata cacctac

37

<210> 46
 <211> 31
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer

<400> 46
 gcgcggcgc tcactcctcc tctgagctgt c

31

<210> 47
 <211> 5
 <212> PRT
 <213> Unknown Organism

<220>
 <223> Description of Unknown Organism: LFCSE motif

<400> 47
 Leu Phe Cys Ser Glu
 1 5

<210> 48
 <211> 2151
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (70)..(1779)

<400> 48
 tgcgttggt gggcctggc accaaagggg cgccccggc ggagagcgga cccagtggcc 60

tggcgatt atg gac ccg gcc gag gcg gtg ctg caa gag aag gca ctc aag 111
 Met Asp Pro Ala Glu Ala Val Leu Gln Glu Lys Ala Leu Lys
 1 5 10

ttt atg aat tcc tca gag aga gaa gac tgt aat aat ggc gaa ccc cct 159
 Phe Met Asn Ser Ser Glu Arg Glu Asp Cys Asn Asn Gly Glu Pro Pro
 15 20 25 30

agg aag ata ata cca gag aag aat tca ctt aga cag aca tac aac agc 207
 Arg Lys Ile Ile Pro Glu Lys Asn Ser Leu Arg Gln Thr Tyr Asn Ser
 35 40 45

tgt gcc aga ctc tgc tta aac caa gaa aca gta tgt tta gca agc act 255
 Cys Ala Arg Leu Cys Leu Asn Gln Glu Thr Val Cys Leu Ala Ser Thr
 50 55 60

gct atg aag act gag aat tgt gtg gcc aaa aca aaa ctt gcc aat ggc 303
 Ala Met Lys Thr Glu Asn Cys Val Ala Lys Thr Lys Leu Ala Asn Gly
 65 70 75

act tcc agt atg att gtg ccc aag caa cgg aaa ctc tca gca agc tat 351
 Thr Ser Ser Met Ile Val Pro Lys Gln Arg Lys Leu Ser Ala Ser Tyr
 80 85 90

gaa aag gaa aag gaa ctg tgt gtc aaa tac ttt gag cag tgg tca gag 399
 Glu Lys Glu Lys Glu Leu Cys Val Lys Tyr Phe Glu Gln Trp Ser Glu
 95 100 105 110

tca gat caa gtg gaa ttt gtg gaa cat ctt ata tcc caa atg tgt cat 447
 Ser Asp Gln Val Glu Phe Val Glu His Leu Ile Ser Gln Met Cys His
 115 120 125

tac caa cat ggg cac ata aac tcg tat ctt aaa cct atg ttg cag aga 495
 Tyr Gln His Gly His Ile Asn Ser Tyr Leu Lys Pro Met Leu Gln Arg
 130 135 140

gat ttc ata act gct ctg cca gct cgg gga ttg gat cat atc gct gag 543
 Asp Phe Ile Thr Ala Leu Pro Ala Arg Gly Leu Asp His Ile Ala Glu
 145 150 155

aac att ctg tca tac ctg gat gcc aaa tca cta tgt gct gct gaa ctt 591
 Asn Ile Leu Ser Tyr Leu Asp Ala Lys Ser Leu Cys Ala Ala Glu Leu
 160 165 170

gtg tgc aag gaa tgg tac cga gtg acc tct gat ggc atg ctg tgg aag 639
 Val Cys Lys Glu Trp Tyr Arg Val Thr Ser Asp Gly Met Leu Trp Lys
 175 180 185 190

aag ctt atc gag aga atg gtc agg aca gat tct ctg tgg aga ggc ctg 687
 Lys Leu Ile Glu Arg Met Val Arg Thr Asp Ser Leu Trp Arg Gly Leu
 195 200 205

gca gaa cga aga gga tgg gga cag tat tta ttc aaa aac aaa cct cct 735
 Ala Glu Arg Arg Gly Trp Gly Gln Tyr Leu Phe Lys Asn Lys Pro Pro
 210 215 220

gac ggg aat gct cct ccc aac tct ttt tat aga gca ctt tat cct aaa	783
Asp Gly Asn Ala Pro Pro Asn Ser Phe Tyr Arg Ala Leu Tyr Pro Lys	
225 230 235	
att ata caa gac att gag aca ata gaa tct aat tgg aga tgt gga aga	831
Ile Ile Gln Asp Ile Glu Thr Ile Glu Ser Asn Trp Arg Cys Gly Arg	
240 245 250	
cat agt tta cag aga att cac tgc cga agt gaa aca agc aaa gga gtt	879
His Ser Leu Gln Arg Ile His Cys Arg Ser Glu Thr Ser Lys Gly Val	
255 260 265 270	
tac tgt tta cag tat gat gat cag aaa ata gta agc ggc ctt cga gac	927
Tyr Cys Leu Gln Tyr Asp Asp Gln Lys Ile Val Ser Gly Leu Arg Asp	
275 280 285	
aac aca atc aag atc tgg gat aaa aac aca ttg gaa tgc aag cga att	975
Asn Thr Ile Lys Ile Trp Asp Lys Asn Thr Leu Glu Cys Lys Arg Ile	
290 295 300	
ctc aca ggc cat aca ggt tca gtc ctc tgt ctc cag tat gat gag aga	1023
Leu Thr Gly His Thr Gly Ser Val Leu Cys Leu Gln Tyr Asp Glu Arg	
305 310 315	
gtg atc ata aca gga tca tgc gat tcc acg gtc aga gtg tgg gat gta	1071
Val Ile Ile Thr Gly Ser Ser Asp Ser Thr Val Arg Val Trp Asp Val	
320 325 330	
aat aca ggt gaa atg cta aac acg ttg att cac cat tgt gaa gca gtt	1119
Asn Thr Gly Glu Met Leu Asn Thr Leu Ile His His Cys Glu Ala Val	
335 340 345 350	
ctg cac ttg cgt ttc aat aat ggc atg atg gtg acc tgc tcc aaa gat	1167
Leu His Leu Arg Phe Asn Asn Gly Met Met Val Thr Cys Ser Lys Asp	
355 360 365	
cgt tcc att gct gta tgg gat atg gcc tcc cca act gac att acc ctc	1215
Arg Ser Ile Ala Val Trp Asp Met Ala Ser Pro Thr Asp Ile Thr Leu	
370 375 380	
cgg agg gtg ctg gtc gga cac cga gct gct gtc aat gtt gta gac ttt	1263
Arg Arg Val Leu Val Gly His Arg Ala Ala Val Asn Val Val Asp Phe	
385 390 395	
gat gac aag tac att gtt tct gca tct ggg gat aga act ata aag gta	1311
Asp Asp Lys Tyr Ile Val Ser Ala Ser Gly Asp Arg Thr Ile Lys Val	
400 405 410	
tgg aac aca agt act tgt gaa ttt gta agg acc tta aat gga cac aaa	1359
Trp Asn Thr Ser Thr Cys Glu Phe Val Arg Thr Leu Asn Gly His Lys	
415 420 425 430	
cga ggc att gcc tgt ttg cag tac agg gac agg ctg gta gtg agt ggc	1407
Arg Gly Ile Ala Cys Leu Gln Tyr Arg Asp Arg Leu Val Val Ser Gly	
435 440 445	
tca tct gac aac act atc aga tta tgg gac ata gaa tgt ggt gca tgt	1455
Ser Ser Asp Asn Thr Ile Arg Leu Trp Asp Ile Glu Cys Gly Ala Cys	
450 455 460	

Ile Ile Pro Glu Lys Asn Ser Leu Arg Gln Thr Tyr Asn Ser Cys Ala
35 40 45

Arg Leu Cys Leu Asn Gln Glu Thr Val Cys Leu Ala Ser Thr Ala Met
 50 55 60

Lys Thr Glu Asn Cys Val Ala Lys Thr Lys Leu Ala Asn Gly Thr Ser
 65 70 75 80

Ser Met Ile Val Pro Lys Gln Arg Lys Leu Ser Ala Ser Tyr Glu Lys
 85 90 95

Glu Lys Glu Leu Cys Val Lys Tyr Phe Glu Gln Trp Ser Glu Ser Asp
 100 105 110

Gln Val Glu Phe Val Glu His Leu Ile Ser Gln Met Cys His Tyr Gln
 115 120 125

His Gly His Ile Asn Ser Tyr Leu Lys Pro Met Leu Gln Arg Asp Phe
 130 135 140

Ile Thr Ala Leu Pro Ala Arg Gly Leu Asp His Ile Ala Glu Asn Ile
 145 150 155 160

Leu Ser Tyr Leu Asp Ala Lys Ser Leu Cys Ala Ala Glu Leu Val Cys
 165 170 175

Lys Glu Trp Tyr Arg Val Thr Ser Asp Gly Met Leu Trp Lys Lys Leu
 180 185 190

Ile Glu Arg Met Val Arg Thr Asp Ser Leu Trp Arg Gly Leu Ala Glu
 195 200 205

Arg Arg Gly Trp Gly Gln Tyr Leu Phe Lys Asn Lys Pro Pro Asp Gly
 210 215 220

Asn Ala Pro Pro Asn Ser Phe Tyr Arg Ala Leu Tyr Pro Lys Ile Ile
 225 230 235 240

Gln Asp Ile Glu Thr Ile Glu Ser Asn Trp Arg Cys Gly Arg His Ser
 245 250 255

Leu Gln Arg Ile His Cys Arg Ser Glu Thr Ser Lys Gly Val Tyr Cys
 260 265 270

Leu Gln Tyr Asp Asp Gln Lys Ile Val Ser Gly Leu Arg Asp Asn Thr
 275 280 285

Ile Lys Ile Trp Asp Lys Asn Thr Leu Glu Cys Lys Arg Ile Leu Thr
 290 295 300

Gly His Thr Gly Ser Val Leu Cys Leu Gln Tyr Asp Glu Arg Val Ile
 305 310 315 320

Ile Thr Gly Ser Ser Asp Ser Thr Val Arg Val Trp Asp Val Asn Thr
 325 330 335

Gly Glu Met Leu Asn Thr Leu Ile His His Cys Glu Ala Val Leu His
 340 345 350

Leu Arg Phe Asn Asn Gly Met Met Val Thr Cys Ser Lys Asp Arg Ser
 355 360 365

Ile Ala Val Trp Asp Met Ala Ser Pro Thr Asp Ile Thr Leu Arg Arg
 370 375 380

Val Leu Val Gly His Arg Ala Ala Val Asn Val Val Asp Phe Asp Asp
 385 390 395 400

Lys Tyr Ile Val Ser Ala Ser Gly Asp Arg Thr Ile Lys Val Trp Asn
 405 410 415

Thr Ser Thr Cys Glu Phe Val Arg Thr Leu Asn Gly His Lys Arg Gly
 420 425 430

Ile Ala Cys Leu Gln Tyr Arg Asp Arg Leu Val Val Ser Gly Ser Ser
 435 440 445

Asp Asn Thr Ile Arg Leu Trp Asp Ile Glu Cys Gly Ala Cys Leu Arg
 450 455 460

Val Leu Glu Gly His Glu Glu Leu Val Arg Cys Ile Arg Phe Asp Asn
 465 470 475 480

Lys Arg Ile Val Ser Gly Ala Tyr Asp Gly Lys Ile Lys Val Trp Asp
 485 490 495

Leu Val Ala Ala Leu Asp Pro Arg Ala Pro Ala Gly Thr Leu Cys Leu
 500 505 510

Arg Thr Leu Val Glu His Ser Gly Arg Val Phe Arg Leu Gln Phe Asp
 515 520 525

Glu Phe Gln Ile Val Ser Ser Ser His Asp Asp Thr Ile Leu Ile Trp
530 535 540

Asp Phe Leu Asn Asp Pro Ala Ala Gln Ala Glu Pro Pro Arg Ser Pro
545 550 555 560

Ser Arg Thr Tyr Thr Tyr Ile Ser Arg
565